Graduate Texts in Mathematics

Béla Bollobás

Graph Theory

An Introductory Course

Springer-Verlag
New York Heidelberg Berlin

Graduate Texts in Mathematics

63

Béla Bollobás

Graph Theory
An Introductory Course

Springer-Verlag
New York Heidelberg Berlin

Béla Bollobás
Department of Pure Mathematics
and Mathematical Statistics
University of Cambridge
16 Mill Lane
Cambridge CB2 1SB
ENGLAND

AMS Subject Classification: 05Cxx

With 80 Figures

Library of Congress Cataloging in Publication Data

Bollobás, Béla.
Graph theory.

(Graduate texts in mathematics: 63)
Includes index.
1. Graph theory. I. Title. II. Series.
QA166.B662 511'.5 79-10720

Softcover reprint of the hardcover 1st edition 1979

9 8 7 6 5 4 3 2 1

ISBN-13: 978-1-4612-9969-1 e-ISBN-13: 978-1-4612-9967-7
DOI: 10.1007?978-1-4612-9967-7

To Gabriella

There is no permanent place in the world for ugly mathematics.

G. H. Hardy
A Mathematician's Apology

Preface

This book is intended for the young student who is interested in graph theory and wishes to study it as part of his mathematical education. Experience at Cambridge shows that none of the currently available texts meet this need. Either they are too specialized for their audience or they lack the depth and development needed to reveal the nature of the subject.

We start from the premise that graph theory is one of several courses which compete for the student's attention and should contribute to his appreciation of mathematics as a whole. Therefore, the book does not consist merely of a catalogue of results but also contains extensive descriptive passages designed to convey the flavour of the subject and to arouse the student's interest. Those theorems which are vital to the development are stated clearly, together with full and detailed proofs. The book thereby offers a leisurely introduction to graph theory which culminates in a thorough grounding in most aspects of the subject.

Each chapter contains three or four sections, exercises and bibliographical notes. Elementary exercises are marked with a $^-$ sign, while the difficult ones, marked by $^+$ signs, are often accompanied by detailed hints. In the opening sections the reader is led gently through the material: the results are rather simple and their easy proofs are presented in detail. The later sections are for those whose interest in the topic has been excited: the theorems tend to be deeper and their proofs, which may not be simple, are described more rapidly. Throughout this book the reader will discover connections with various other branches of mathematics, including optimization theory, linear algebra, group theory, projective geometry, representation theory, probability theory, analysis, knot theory and ring theory. Although most of these connections are not essential for an understanding of the book, the reader would benefit greatly from a modest acquaintance with these subjects.

The bibliographical notes are not intended to be exhaustive but rather to guide the reader to additional material.

I am grateful to Andrew Thomason for reading the manuscript carefully and making many useful suggestions. John Conway has also taught the graph theory course at Cambridge and I am particularly indebted to him for detailed advice and assistance with Chapters II and VIII. I would like to thank Springer-Verlag and especially Joyce Schanbacher for their efficiency and great skill in producing this book.

Cambridge
April 1979

Béla Bollobás

Contents

CHAPTER I
Fundamentals

The purpose of this introduction is to familiarise the reader with the basic concepts and results of graph theory. The chapter inevitably contains a large number of definitions and in order to prevent the reader growing weary we prove simple results as soon as possible. The reader is not expected to have complete mastery of Chapter I before sampling the rest of the book, indeed, he is encouraged to skip ahead since most of the terminology is self-explanatory. We should add at this stage that the terminology of graph theory is far from being standard, though that used in this book is well accepted.

§1 Definitions

A *graph* G is an ordered pair of disjoint sets (V, E) such that E is a subset of the set of unordered pairs of V. Unless it is explicitly stated otherwise, we consider only finite graphs, that is V and E are always finite. The set V is the set of *vertices* and E is the set of *edges*. If G is a graph then $V = V(G)$ is the vertex set of G and $E = E(G)$ is the edge set. An edge $\{x, y\}$ is said to *join* the vertices x and y and is denoted by xy. Thus xy and yx mean exactly the same edge; the vertices x and y are the *endvertices* of this edge. If $xy \in E(G)$ then x and y are *adjacent* or *neighbouring* vertices of G and the vertices x and y are *incident* with the edge xy. Two edges are *adjacent* if they have exactly one common endvertex.

As the terminology suggests, we do not usually think of a graph as an ordered pair, but as a collection of vertices some of which are joined by

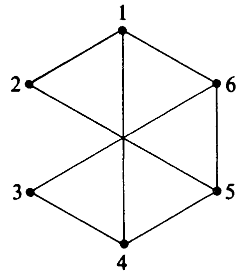


Figure I.1. A graph.

edges. It is then a natural step to draw a picture of the graph. In fact, sometimes the easiest way to describe a graph is to draw it; the graph $G = (\{1,2,3,4,5,6\}, \{12,14,16,25,34,36,45,56\})$ is immediately comprehended by looking at Figure I.1.

We say that $G' = (V', E')$ is a *subgraph* of $G = (V, E)$ if $V' \subset V$ and $E' \subset E$. In this case we write $G' \subset G$. If G' contains *all edges* of G that join two vertices in V' then G' is said to be the subgraph *induced* or *spanned by* V' and is denoted by $G[V']$. A subgraph G' of G is an *induced* subgraph if $G' = G[V(G')]$. These concepts are illustrated in Figure I.2.

We shall often construct new graphs from old ones by deleting or adding some vertices and edges. If $W \subset V(G)$ then $G - W = G[V\backslash W]$ is the subgraph of G obtained by deleting the vertices in W *and all edges incident with them.* Similarly if $E' \subset E(G)$ then $G - E' = (V(G), E(G)\backslash E')$. If $W = \{w\}$ and $E' = \{xy\}$ then this notation is simplified to $G - w$ and $G - xy$. Similarly, if x and y are non-adjacent vertices of G then $G + xy$ is obtained from G by joining x to y.

If x is a vertex of a graph G then instead of $x \in V(G)$ we usually write $x \in G$. The *order* of G is the number of vertices; it is denoted by $|G|$. The same notation is used for the number of elements (cardinality) of a *set*: $|X|$ denotes the number of elements of the set X. Thus $|G| = |V(G)|$. The *size* of G is the number of edges; it is denoted by $e(G)$. We write G^n for an *arbitrary graph of order n.* Similarly $G(n, m)$ denotes an *arbitrary graph of order n and size m.*

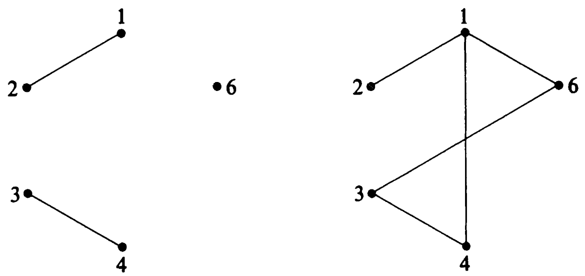


Figure I.2. A subgraph and an induced subgraph of the graph in Figure I.1.

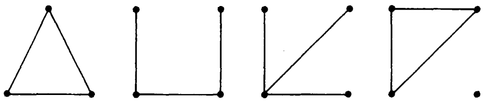


Figure I.3. Graphs of order at most 4 and size 3.

Two graphs are *isomorphic* if there is a correspondence between their vertex sets that preserves adjacency. Thus $G = (V, E)$ is isomorphic to $G' = (V', E')$ if there is a bijection $\phi: V \to V'$ such that $xy \in E$ iff $\phi(x)\phi(y) \in E'$. Clearly isomorphic graphs have the same order and size. Usually we do not distinguish between isomorphic graphs, unless we consider graphs with a distinguished or labelled set of vertices (for example, subgraphs of a given graph). In accordance with this convention, if G and H are isomorphic graphs then we write either $G \cong H$ or simply $G = H$. In Figure I.3 we show all graphs (within isomorphism) that have order at most 4 and size 3.

The size of a graph of order n is at least 0 and at most $\binom{n}{2}$. Clearly for every m, $0 \leq m \leq \binom{n}{2}$, there is a graph $G(n, m)$. A graph of order n and size $\binom{n}{2}$ is called a *complete n-graph* and is denoted by K^n; an *empty n-graph* E^n has order n and no edges. In K^n every two vertices are adjacent, while in E^n no two vertices are adjacent. The graph $K^1 = E^1$ is said to be *trivial.*

The set of vertices adjacent to a vertex $x \in G$ is denoted by $\Gamma(x)$. The *degree* of x is $d(x) = |\Gamma(x)|$. If we want to emphasize that the underlying graph is G then we write $\Gamma_G(x)$ and $d_G(x)$; a similar convention will be adopted for other functions depending on an underlying graph. Thus if $x \in H = G[W]$ then

$$\Gamma_H(x) = \{y \in H: xy \in E(H)\} = \Gamma_G(x) \cap W.$$

The *minimum degree* of the vertices of a graph G is denoted by $\delta(G)$ and the *maximum degree* by $\Delta(G)$. A vertex of degree 0 is said to be an *isolated* vertex. If $\delta(G) = \Delta(G) = k$, that is every vertex of G has degree k then G is said to be *k-regular* or *regular of degree k.* A graph is *regular* if it is k-regular for some k. A 3-regular graph is said to be *cubic.*

If $V(G) = \{x_1, x_2, \ldots, x_n\}$ then $(d(x_i))_1^n$ is a *degree sequence* of G. Usually we order the vertices in such a way that the degree sequence obtained in this way is monotone increasing or monotone decreasing, for example $\delta(G) = d(x_1) \leq \cdots \leq d(x_n) = \Delta(G)$. Since each edge has two endvertices, the sum of the degrees is exactly twice the number of edges:

$$\sum_1^n d(x_i) = 2e(G). \tag{1}$$

In particular, the sum of degrees is even:

$$\sum_1^n d(x_i) \equiv 0 \pmod 2. \tag{2}$$

This last observation is sometimes called the *handshaking lemma*, since it expresses the fact that in any party the total number of hands shaken is even. Equivalently, (2) states that the number of vertices of *odd* degree is *even*. We see also from (1) that $\delta(G) \le \lfloor 2e(G)/n \rfloor$ and $\Delta(G) \ge \lceil 2e(G)/n \rceil$. Here $\lfloor x \rfloor$ denotes the greatest integer not greater than x and $\lceil x \rceil = -\lfloor -x \rfloor$.

A *path* is a graph P of the form

$$V(P) = \{x_0, x_1, \dots, x_l\}, \qquad E(P) = \{x_0x_1, x_1x_2, \dots, x_{l-1}x_l\}.$$

This path P is usually denoted by $x_0x_1 \dots x_l$. The vertices x_0 and x_l are the *endvertices* of P and $l = e(P)$ is the *length* of P. We say that P is a path *from* x_0 *to* x_l or an x_0-x_l *path*. Of course, P is also a path from x_l to x_0 or an x_l-x_0 path. Sometimes we wish to emphasize that P is considered to go from x_0 to x_l and then call x_0 the *initial* and x_l the *terminal* vertex of P. A path with initial vertex x is an x-*path*.

The term *independent* will be used in connection with vertices, edges and paths of a graph. A set of *vertices* (*edges*) is *independent* if no two elements of it are adjacent; a set of paths is *independent* if for any two paths each vertex belonging to both paths is an endvertex of both. Thus $P_1, P_2, \dots, P_k$ are independent x-y paths iff $V(P_i) \cap V(P_j) = \{x, y\}$ whenever $1 \neq j$. Also, $W \subset V(G)$ consists of independent vertices iff $G[W]$ is an empty graph.

Most paths we consider are subgraphs of a given graph G. A *walk* W in G is an alternating sequence of vertices and edges, say $x_0, \alpha_1, x_2, \alpha_2, \dots, \alpha_l, x_0$ where $\alpha_i = x_{i-1}x_i$, $0 \le i < l$. In accordance with the terminology above, W is an x_0-x_l *walk* and is denoted by $x_0x_1 \dots x_l$; the *length* of W is l. This walk W is called a *trail* if all its edges are distinct. Note that a path is a walk with distinct vertices. A trail whose endvertices coincide (a *closed* trail) is called a *circuit*. If a walk $W = x_0x_1 \dots x_l$ is such that $l \ge 3$, $x_0 = x_l$ and the vertices x_i, $0 < i < l$, are distinct from each other and x_0 then W is said to be a *cycle*. For simplicity this cycle is denoted by $x_1x_2 \dots x_l$. Note that the notation differs from that of a path since x_1x_l is also an edge of this cycle. Furthermore, $x_1x_2 \dots x_l$, $x_lx_{l-1} \dots x_1$, $x_2x_3 \dots x_lx_1$, $x_ix_{i-1} \dots x_1x_lx_{l-1} \dots x_{i+1}$ all denote the same cycle.

The symbol P^l denotes an arbitrary *path of length* l and C^l denotes a *cycle of length* l. We call C^3 a triangle, C^4 a quadrilateral, C^5 a pentagon, etc. (See Figure I.4). A cycle is *even* (*odd*) if its length is even (odd).

Given vertices x, y, their *distance* $d(x, y)$ is the minimum length of an x-y path. If there is no x-y path then $d(x, y) = \infty$.

A graph is *connected* if for every pair $\{x, y\}$ of distinct vertices there is a path from x to y. Note that a connected graph of order at least 2 cannot contain an isolated vertex. A *maximal connected subgraph* is a *component*

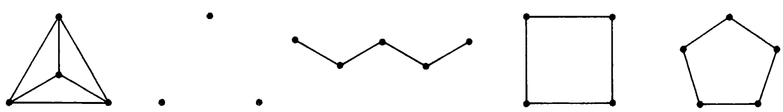


Figure I.4. The graphs K^4, E^3, P^4, C^4 and C^5.

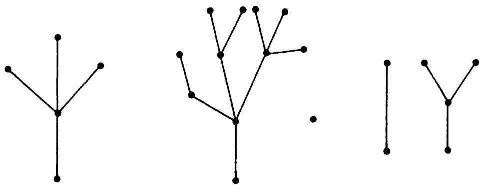


Figure I.5. A forest.

of the graph. A *cutvertex* is a vertex whose deletion increases the number of components. Similarly an edge is a *bridge* if its deletion increases the number of components. Thus an edge of a connected graph is a bridge if its deletion disconnects the graph. A graph without any cycles is a *forest* or an *acyclic* graph; a *tree* is a connected forest. (See Figure I.5.) The relation of a tree to a forest sounds less absurd if we note that a forest is a disjoint union of trees; in other words, a forest is a graph whose every component is a tree.

A graph G is a *bipartite graph with vertex classes* V_1 and V_2 if $V(G) = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$ and each edge joins a vertex of V_1 to a vertex of V_2. Similarly G is *r-partite with vertex classes* $V_1, V_2, \ldots, V_r$ if $V(G) = V_1 \cup V_2 \cup \cdots \cup V_r$, $V_i \cap V_j = \emptyset$ whenever $1 \le i < j \le r$, and no edge joins two vertices in the same class. The graphs in Figure I.1 and Figure I.5 are bipartite. The symbol $K(n_1, \ldots, n_r)$ denotes a *complete r-partite graph*: it has n_i vertices in the ith class and contains all edges joining vertices in distinct classes. For simplicity we often write $K^{p,q}$ instead of $K(p, q)$ and $K_r(t)$ instead of $K(t, \ldots, t)$.

We shall write $G \cup H = (V(G) \cup V(H), E(G) \cup E(H))$ and kG for the union of k disjoint copies of G. We obtain the *join* $G + H$ from $G \cup H$ by adding all edges between G and H. Thus, for example, $K^{2,3} = E^2 + E^3$ and $K_r(t) = E^t + \cdots + E^t$.

There are several notions closely related to that of a graph. A *hypergraph* is a pair (V, E) such that $V \cap E = \emptyset$ and E is a subset of $\mathscr{P}(V)$, the power set of V, that is the set of all subsets of V. In fact, there is a simple 1-1 correspondence between the class of hypergraphs and the class of certain bipartite graphs. Indeed, given a hypergraph (V, E), construct a bipartite graph with vertex classes V and E by joining a vertex $x \in V$ to a hyperedge $S \in E$ iff $x \in S$.

By definition a graph does not contain a *loop*, an "edge" joining a vertex to itself; neither does it contain *multiple edges*, that is several "edges" joining the same two vertices. In a *multigraph* both multiple edges and multiple loops are allowed; a loop is a special edge.

If the edges are *ordered* pairs of vertices then we get the notions of a *directed graph* and *directed multigraph*. An ordered pair (a, b) is said to be an edge *directed from a to b*. or an edge *beginning at a and ending at b*, and is denoted by $\overrightarrow{ab}$ or simply ab. The notions defined for graphs are easily carried over to multigraphs, directed graphs and directed multigraphs, mutatis mutandis. Thus a (directed) trail in a directed multigraph is an

alternating sequence of vertices and edges: $x_0, e_1, x_1, e_2, \ldots, e_l, x_l$, such that e_i begins at x_{i-1} and ends at x_i.

An *oriented graph* is a directed graph obtained by orienting the edges, that is by giving the edge ab a direction $\vec{ab}$ or $\vec{ba}$. Thus an oriented graph is a directed graph in which at most one of $\vec{ab}$ and $\vec{ba}$ occurs.

§2 Paths, Cycles and Trees

With the concepts defined so far we can start proving some results about graphs. Though these results are hardly more than simple observations, and our main aim in presenting them is to familiarize the reader with the concepts, in keeping with the style of the other chapters we shall call them theorems.

Theorem 1. *Let x be a vertex of a graph G and let W be the vertex set of a component containing x. Then the following assertions hold.*

i. $W = \{y \in G\colon G \text{ contains an } x\text{-}y \text{ path}\}$.
ii. $W = \{y \in G\colon G \text{ contains an } x\text{-}y \text{ trail}\}$.
iii. $W = \{y \in G\colon d(x, y) < \infty\}$.
iv. *For $u, v \in V = V(G)$ put uRv iff $uv \in E(G)$, and let $\tilde{R}$ be the smallest equivalence relation on V containing R. Then W is the equivalence class of x.* □

This little result implies that every graph is the vertex disjoint union of its components (equivalently, every vertex is contained in a unique component), and an edge is a bridge iff it is not contained in a cycle.

Theorem 2. *A graph is bipartite iff it does not contain an odd cycle.*

Proof. Suppose G is bipartite with vertex classes V_1 and V_2. Let $x_1 x_2 \ldots x_l$ be a cycle in G. We may assume that $x_1 \in V_1$. Then $x_2 \in V_2$, $x_3 \in V_1$, and so on: $x_i \in V_1$ iff i is odd. Since $x_l \in V_2$, we find that l is even.

Suppose now that G does not contain an odd cycle. Since a graph is bipartite iff each component of it is, we may assume that G is connected. Pick a vertex $x \in V(G)$ and put $V_1 = \{y\colon d(x, y) \text{ is odd}\}$. There is no edge joining two vertices of the same class since otherwise G would contain an odd cycle. Hence G is bipartite. □

Theorem 3. *A graph is a forest iff for every pair $\{x, y\}$ of distinct vertices it contains at most one x-y path.*

Proof. If $x_1 x_2 \ldots x_l$ is a cycle in a graph G then $x_1 x_2 \ldots x_l$ and $x_1 x_l$ are two x_1-x_l paths in G.

Conversely, let $P_1 = x_0x_1 \dots x_l$ and $P_2 = x_0y_1y_2 \dots y_kx_l$ be two distinct x_0-x_l paths in a graph G. Let $i + 1$ be the minimal index for which $x_{i+1} \neq y_{i+1}$, and let j be the minimal index for which $j \geq i$ and y_{j+1} is a vertex of P_1, say $y_{j+1} = x_h$. Then $x_ix_{i+1} \dots x_hy_jy_{j-1} \dots y_{i+1}$ is a cycle in G. □

Theorem 4. *The following assertions are equivalent for a graph G.*

i. *G is a tree.*
ii. *G is a minimal connected graph, that is G is connected and if $xy \in E(G)$ then $G - xy$ is disconnected. [In other words, G is connected and every edge is a bridge.]*
iii. *G is a maximal acyclic graph, that is G is acyclic and if x and y are non-adjacent vertices of G then $G + xy$ contains a cycle.*

PROOF. Suppose G is a tree. Let $xy \in E(G)$. The graph $G - xy$ cannot contain an x-y *path* $xz_1z_2 \dots z_ky$ since otherwise G contains the *cycle* $xz_1z_2 \dots z_ky$. Hence $G - xy$ is disconnected and so G is a minimal connected graph. Similarly if x and y are non-adjacent vertices of the tree G then G contains a *path* $xz_1z_2 \dots z_ky$ and so $G + xy$ contains the *cycle* $xz_1z_2 \dots z_ky$. Hence $G + xy$ contains a cycle and so G is a maximal acyclic graph.

Suppose next that G is a minimal connected graph. If G contains a *cycle* $xz_1z_2 \dots z_ky$ then $G - xy$ is still connected since in any $u - v$ walk in G the edge xy can be replaced by the *path* $xz_1z_2 \dots x_ky$. As this contradicts the minimality of G, we conclude that G is acyclic so it is a tree.

Suppose finally that G is a maximal acyclic graph. Is G connected? Yes, since if x and y belong to different components, the addition of xy to G cannot create a *cycle* $xz_1z_2 \dots z_ky$ since otherwise the path $xz_1z_2 \dots z_ky$ is in G. Thus G is a tree. □

Corollary 5. *Every connected graph contains a spanning tree, that is a tree containing every vertex of the graph.*

PROOF. Take a minimal connected spanning subgraph. There are several simple constructions of a spanning tree of a graph G; we present two of them. Pick a vertex x and put $V_i = \{y \in G: d(x, y) = i\}$, $i = 0, 1, \dots$. Note that if $y_i \in V_i$, $i > 0$, and $xz_1z_2 \dots z_{i-1}y_i$ is an $x - y_i$ path (whose existence is guaranteed by the definition of V_i) then $d(x, z_j) = j$ for every j, $0 < j < i$. In particular, $V_j \neq \emptyset$ for $0 \leq j \leq i$ and for every $y \in V_i$, $i > 0$, there is a vertex $y' \in V_{i-1}$ joined to y. (Of course, this vertex y' is usually not unique but for each $y \neq x$ we *pick* only one y'.) Let T be the subgraph of G with vertex set V and edge set $E(T) = \{yy': y \neq x\}$. Then T is connected since every $y \in V - \{x\}$ is joined to x by a path $yy'y'' \dots x$. Furthermore, T is acyclic since if W is any subset of V and w is a vertex in W furthest from x then w is joined to at most one vertex in W. Thus T is a spanning tree. □

The argument above shows that with $k = \max_y d(x, y)$ we have $V_i \neq \emptyset$ for $0 \leq i \leq k$ and $V = V(G) = \bigcup_0^k V_i$. At this point it is difficult to resist the remark that $\text{diam } G = \max_{x,y} d(x, y)$ is called the *diameter* of G and $\text{rad } G = \min_x \max_y d(x, y)$ is the *radius* of G.

If we choose $x \in G$ with $k = \max_y d(x, y) = \text{rad } G$ then the spanning tree T also has radius k.

A slight variant of the above construction of T goes as follows. Pick $x \in G$ and let T_1 be the subgraph of G with this single vertex x. Then T_1 is a tree. Suppose we have constructed trees $T_1 \subset T_2 \subset \cdots \subset T_k \subset G$, where T_i has order i. If $k < n = |G|$ then by the connectedness of G there is a vertex $y \in V(G) - V(T_k)$ that is adjacent (in G) to a vertex $z \in T_k$. Let T_{k+1} be obtained from T_k by adding to it the vertex y and the edge yz. Then T_{k+1} is connected and as yz cannot be an edge of a cycle in T_{k+1}, it is also acyclic. Thus T_{k+1} is also a tree, so the sequence $T_0 \subset T_1 \subset \ldots$ can be continued to T_n. This tree T_n is then a spanning tree of G.

The spanning trees constructed by either of the methods above have order n (of course!) and size $n - 1$. In the first construction there is a 1-1 correspondence between $V - \{x\}$ and $E(T)$, given by $y \to yy'$, and in the second construction $e(T_k) = k - 1$ for each k since $e(T_1) = 0$ and T_{k+1} has one more edge than T_k. Since by Theorem 4 every tree has a *unique* spanning tree, namely itself, we have arrived at the following result.

Corollary 6. *A tree of order n has size $n - 1$; a forest of order n with k components has size $n - k$.* □

The first part of this corollary can be incorporated into several other characterizations of trees. In particular, a graph of order n is a tree iff it is connected and has size $n - 1$. The reader is invited to prove these characterizations (Exercise 9).

Corollary 7. *A tree of order at least 2 contains at least 2 vertices of degree 1.*

Proof. Let $d_1 \leq d_2 \leq \cdots \leq d_n$ be the degree sequence of a tree T of order $n \geq 2$. Since T is connected, $\delta(T) = d_1 \geq 1$. Hence if T had at most one vertex of degree 1, by (1) and Corollary 5 we would have

$$2e(T) = 2n - 2 = \sum_1^n d_i \geq 1 + 2(n - 1). \qquad \square$$

A well known problem in optimization theory asks for a relatively easy way of finding a spanning subgraph with a special property. Given a graph $G = (V, E)$ and a positive valued cost function f defined on the edges, $f\colon E \to \mathbb{R}^+$, find a connected spanning subgraph $T = (V, E')$ of G for which

$$f(T) = \sum_{xy \in E'} f(xy)$$

is minimal. We call such a spanning subgraph T an *economical* spanning subgraph. One does not need much imagination to translate this into a "real life" problem. Suppose certain villages in an area are to be joined to a water supply situated in one of the villages. The system of pipes is to consist of pipelines connecting the water towers of two villages. For any two villages we know how much it would cost to build a pipeline connecting them, provided such a pipeline can be built at all. How can we find an economical system of pipes?

In order to reduce the second problem to the above problem about graphs, let G be the graph whose vertex set is the set of villages and in which xy is an edge iff it is possible to build a pipeline joining x to y; denote the cost of such a pipeline by $f(xy)$ (see Figure I.6). Then a system of pipes corresponds to a connected spanning subgraph T of G. Since the system has to be economical, T is a minimal connected spanning subgraph of G, that is a *spanning tree* of G.

The connected spanning subgraph T we look for has to be a *minimal* connected subgraph since otherwise we could find an edge α whose deletion would leave T connected and then $T - \alpha$ would be a more economical spanning subgraph. Thus T is a spanning tree of G. Corresponding to the various characterizations and constructions of a spanning tree, we have several easy ways of finding an economical spanning tree; we shall describe four of these methods.

(1) Given G and $f: E(G) \to \mathbb{R}$, we choose one of the cheapest edges of G, that is an edge α for which $f(\alpha)$ is minimal. Each subsequent edge will be chosen from among the cheapest remaining edges of G with the only restriction that we must not select all edges of any cycle, that is the subgraph of G formed by the selected edges is acyclic.

The process terminates when no edge can be added to the set E' of edges selected so far without creating a cycle. Then $T_1 = (V(G), E')$ is a maximal acyclic subgraph of G so, by Theorem 4(iii) it is a spanning tree of G.

(2) This method is based on the fact that it is foolish to use a costly edge unless it is needed to ensure the connectedness of the subgraph. Thus let us delete one by one those costliest edges whose deletion does not disconnect the graph. By Theorem 4(ii) the process ends in a spanning tree T_2.

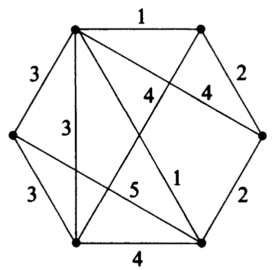


Figure I.6. A graph with a function $f: E \to \mathbb{R}^+$; the number next to an edge xy is the cost $f(xy)$ of the edge.

(3) Pick a vertex x_1 of G and select one of the least costly edges incident with x_1, say x_1x_2. Then choose one of the least costly edges of the form x_ix, where $1 \le i \le 2$ and $x \notin \{x_1, x_2\}$. Having found vertices $x_1, x_2, \ldots, x_k$ and an edge x_ix_j, $i < j$, for each vertex x_j with $j \le k$, select one of the least costly edges of the form x_ix, say x_ix_{k+1} where $1 \le i \le k$ and $x_{k+1} \notin \{x_1x_2, \ldots, x_k\}$. The process terminates after we have selected $n - 1$ edges. Denote by T_3 the spanning tree given by these edges. (See Figure I.7.)

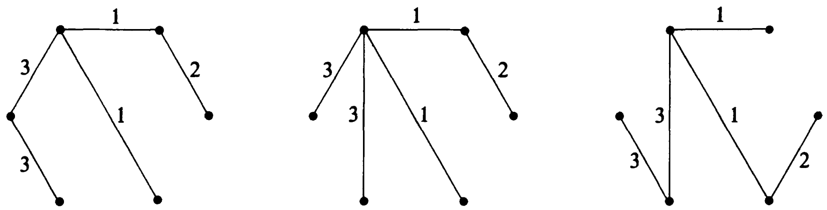


Figure I.7. Three of the six economical spanning trees of the graph shown in Figure I.6.

(4) This method is applicable only if no two pipelines cost the same. The advantage of the method is that every village can make its own decision and start building a pipeline without bothering to find out what the other villages are going to do. Of course, each village will start building the cheapest pipeline ending in the village. It may happen that both village x and village y will build the pipeline xy; in this case they meet in the middle and end up with a single pipeline from x to y. Thus at the end of this stage some villages will be joined by pipelines but the whole system of pipes need not be connected. At the next stage each group of villages joined to each other by pipelines finds the cheapest pipeline going to a village not in the group and begins to build that single pipeline. The same procedure is repeated until a connected system is obtained. Clearly the villages will never build all the pipes of a cycle so the final system of pipes will be a spanning tree (See Figure I.8.)

Theorem 8. *Each of the four methods described above produces an economical spanning tree. If no two edges have the same cost then there is a unique economical spanning tree.*

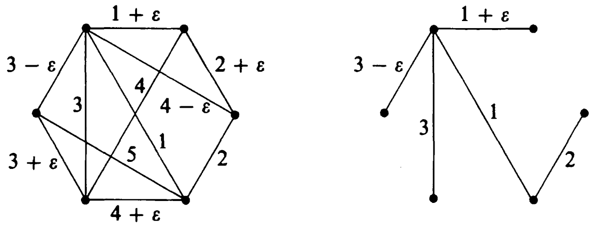


Figure I.8. The graph of Figure I.6 with a slightly altered cost function ($0 < \varepsilon < \frac{1}{2}$) and its unique economical spanning tree.

PROOF. Choose an economical spanning tree T of G that has as many edges in common with T_1 as possible. (T_1 is the spanning tree constructed by the first method.)

Suppose that $E(T_1) \neq E(T)$. The edges of T_1 have been selected one by one; let xy be the first edge of T_1 that is not an edge of T. Then T contains a unique x-y path, say P. This path P has at least one edge, say uv, that does not belong to T_1, since otherwise T_1 would contain a cycle. When xy was selected as an edge of T_1, the edge uv was also a candidate. As xy was chosen and not uv, the edge xy can not be costlier than uv, that is $f(xy) \leq f(uv)$. Then $T' = T - uv + xy$ is a spanning tree and since $f(T') = f(T) - f(uv) + f(xy) \leq f(T)$, T' is an economical spanning tree of G. (Of course, this inequality implies that $f(T') = f(T)$ and $f(xy) = f(uv)$.) This tree T' has more edges in common with T_1 than T, contradicting the choice of T. Hence $T = T_1$ so T_1 is indeed an economical spanning tree.

Slight variants of the proof above show that the spanning trees T_2 and T_3, constructed by the second and third methods, are also economical. We invite the reader to furnish the details (Exercise 19).

Suppose now that no two edges have the same cost, that it $f(xy) \neq f(uv)$ whenever $xy \neq uv$. Let T_4 be the spanning tree constructed by the fourth method and let T be an economical spanning tree. Suppose that $T \neq T_4$, and let xy be the first edge not in T that we select for T_4. The edge xy was selected since it is the *least costly* edge of G joining a vertex of a subtree F of T_4 to a vertex outside F. The x-y path in T has an edge uv joining a vertex of F to a vertex outside F so $f(xy) < f(uv)$. However, this is impossible since $T' = T - uv + xy$ is a spanning tree of G and $f(T') < f(T)$. Hence $T = T_4$. This shows that T_4 is indeed an economical spanning tree. Furthermore, since the spanning tree constructed by the fourth method is unique, the economical spanning tree is unique if no two edges have the same cost. □

§3 Hamilton Cycles and Euler Circuits

The so-called *travelling salesman problem* greatly resembles the economical spanning tree problem discussed in the preceding section, but the similarity is only superficial. A salesman is to make a tour of n cities, at the end of which he has to return to the head office he starts from. The cost of the journey between any two cities is known. The problem asks for an efficient algorithm for finding a least expensive tour. (As we shall not deal with algorithmic problems, we leave the term "efficient" undefined; loosely speaking an algorithm is efficient if the computing time is bounded by a polynomial in the number of vertices.) Though a considerable amount of work has been done on this problem, since its solution would have important practical

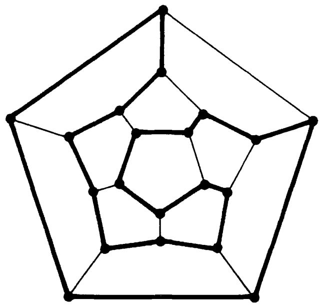


Figure I.9. A Hamiltonian cycle in the graph of the dodecahedron.

applications, it is not known whether or not there is an efficient algorithm for finding a least expensive route.

In another version of the travelling salesman problem the route is required to be a cycle, that is the salesman is not allowed to visit the same city twice (except the city of the head office). A cycle containing all the vertices of a graph is said to be a *Hamilton cycle* of the graph. The origin of this term is a game invented in 1857 by Sir William Rowan Hamilton based on the construction of cycles containing all the vertices in the graph of the dodecahedron (see Figure I.9). A *Hamilton path* of a graph is a path containing all the vertices of the graph. A graph containing a Hamilton cycle is said to be *Hamiltonian.*

In fact, Hamilton cycles and paths in special graphs had been studied well before Hamilton proposed his game. In particular, the puzzle of the knight's tour on a chessboard, thoroughly analysed by Euler in 1759, asks for a Hamilton cycle in the graph whose vertices are the 64 squares of a chessboard and in which two vertices are adjacent if a knight can jump from one square to the other. Figure I.10. shows two solutions of this puzzle.

If in the second, more restrictive version of the travelling salesman problem there are only two travel costs, 1 and ∞ (expressing the impossibility of the journey), then the question is whether or not the graph formed by the edges with travel cost 1 contains a Hamilton cycle. Even this special case of the travelling salesman problem is unsolved: no efficient algorithm is known

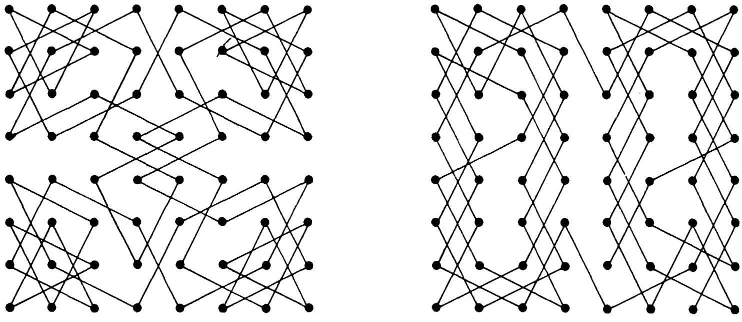


Figure I.10. Two tours of a knight on a chessboard.

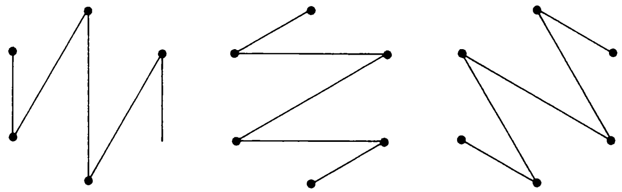


Figure I.11. Three edge-disjoint Hamilton paths in K^6.

for constructing a Hamilton cycle, though neither is it known that there is no such algorithm.

If the travel cost between any two cities is the same, then our salesman has no difficulty in finding a least expensive tour: any permutation of the $n - 1$ cities (the nth city is that of the head office) will do. Revelling in his new found freedom, our salesman decides to connect duty and pleasure and promises not to take the same road xy again whilst there is a road uv he hasn't seen yet. Can he keep his promise? In order to plan a required sequence of journeys for our salesman we have to decompose K^n into the union of some edge disjoint Hamilton cycles. For which values of n is this possible? Since K^n is $(n - 1)$-regular and a Hamilton cycle is 2-regular, a *necessary* condition is that $n - 1$ is even, that is n is odd. This necessary condition also follows from the fact that $e(K^n) = \frac{1}{2}n(n - 1)$ and a Hamilton cycle contains n edges, so K^n has to be the union of $\frac{1}{2}(n - 1)$ Hamilton cycles.

Let us assume now that n is odd, $n \geq 3$. Deleting a vertex of K^n we see that if K^n is the union of $\frac{1}{2}(n - 1)$ Hamilton cycles K^{n-1} is the union of $\frac{1}{2}(n - 1)$ Hamilton paths. (In fact $n - 1$ has to be even if K^{n-1} is the union of some Hamilton paths since $e(K^{n-1}) = \frac{1}{2}(n - 1)(n - 2)$ and a Hamilton path in K^{n-1} has $n - 2$ edges.) With the hint shown in Figure I.11 the reader can show that for odd values of n the graph K^{n-1} is indeed the union of $\frac{1}{2}(n - 1)$ Hamilton paths. In this decomposition of K^{n-1} into $\frac{1}{2}(n - 1)$ Hamilton paths each vertex is the endvertex of exactly one Hamilton path. (In fact, this holds for every decomposition of K^{n-1} into $\frac{1}{2}(n - 1)$ edge-disjoint Hamilton paths since each vertex x of K^{n-1} has odd degree so at least one Hamilton path has to end in x.) Consequently if we add a new vertex to K^{n-1} and complete each Hamilton path in K^{n-1} to a Hamilton cycle in K^n then we obtain a decomposition of K^n into $\frac{1}{2}(n - 1)$ edge disjoint Hamilton cycles. Thus we have proved the following result.

Theorem 9. *For $n \geq 3$ the complete graph K^n is decomposable into edge disjoint Hamilton cycles iff n is odd. For $n \geq 2$ the complete graph K^n is decomposable into edge disjoint Hamilton paths iff n is even.* □

The result above shows that if $n \geq 3$ is odd, then we can string together $\frac{1}{2}(n - 1)$ edge disjoint cycles in K^n to obtain a circuit containing all the edges of K^n. In general a circuit in a graph G containing all the edges is said to be and *Euler circuit* of G. Similarly a trail containing all edges is an *Euler trail.*

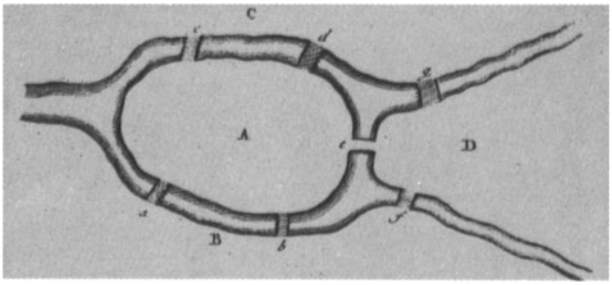


Figure I.12. The seven bridges on the Pregel in Königsberg.

A graph is *Eulerian* if it has an Euler circuit. Euler circuits and trails are named after Leonhard Euler, who in 1736 characterized those graphs which contain them. At the time Euler was a professor of mathematics in St. Petersburg, and was led to the problem by the puzzle of the seven bridges on the Pregel (see Figure I.12) in the ancient Prussian city of Königsberg (birthplace of Kant; since 1944 it has belonged to the USSR and is called Kaliningrad). Could anyone plan a walk in such a way that he would cross each bridge once and only once?

It is clear that such a walk is possible iff the graph in Figure I.13 has an Euler trail.

Theorem 10. *A non-trivial connected graph has an Euler circuit iff each vertex has even degree.*

A connected graph has a Euler trail from a vertex x to a vertex $y \neq x$ iff x and y are the only vertices of odd degree.

Proof. The conditions are clearly necessary. For example, if $x_1 x_2 \ldots x_m$ is an Euler circuit in G and x occurs k times in the sequence $x_1, x_2, \ldots, x_m$, then $d(x) = 2k$.

We prove the sufficiency of the first condition by induction on the number of edges. If there are no edges, there is nothing to prove so we proceed to the induction step.

Let G be a non-trivial connected graph in which each vertex has even degree. Since $e(G) \geq 1$, we find that $\delta(G) \geq 2$ so, by Corollary 7, G contains a cycle. Let C be a *circuit* in G with the maximal number of edges. Suppose

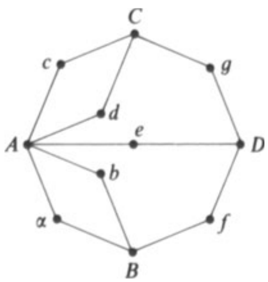


Figure I.13. A graph of the Königsberg bridges.

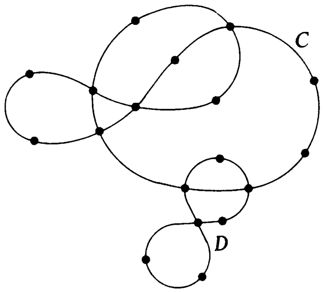


Figure I.14. The circuits C and D.

C is not Eulerian. As G is connected, C contains a vertex x which is in a non-trivial component H of $G - E(C)$. Every vertex of H has even degree in H so by the induction hypothesis H contains an Euler circuit D. The circuits C and D (see Figure I.14) are edge-disjoint and have a vertex in common, so they can be strung together to form a circuit with more edges than C. As this contradicts the maximality of $e(C)$, the circuit C is Eulerian.

Suppose now that G is connected and x and y are the only vertices of odd degree. Let G^* be obtained from G by adding to it a vertex u together with the edges ux and uy. Then by the first part G^* has an Euler circuit C^*. Clearly $C^* - u$ is an Euler trail from x to y. □

Theorem 10 implies that there is no walk satisfying the conditions of the Königsberg bridge puzzle, since the associated graph in Figure I.13 has four vertices of odd degree. The plan of the corridors of an exhibition is also easily turned into a graph: an edge corresponds to a corridor and a vertex to the conjunction of several corridors. If the entrance and exit are the same, a visitor can walk along every corridor exactly once iff the corresponding graph has an Eulerian circuit. In general a visitor must have a plan in order to achieve this: he cannot just walk through any new corridor he happens to come to. However, in a well planned (!) exhibition a visitor would be certain to see all the exhibits provided he avoided going along the same corridor twice and continued his walk until there were no new exhibits ahead of him. The graph of such an exhibition is said to be *randomly Eulerian* from the vertex corresponding to the entrance (which is also the exit). See Figure I.15 for two examples. Randomly Eulerian graphs are also easily characterized (Exercises 24–26).

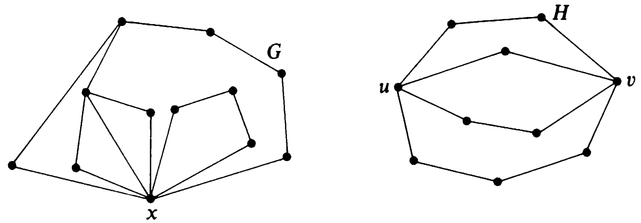


Figure I.15. G is randomly Eulerian from x; H is randomly Eulerian from both u and v.

§4 Planar Graphs

The graph of the corridors of an exhibition is a *planar graph*: it can be drawn in the plane in such a way that no two edges intersect. Putting it a bit more rigorously, it is possible to represent it in the plane in such a way that the vertices correspond to distinct points and the edges to simple Jordan curves connecting the points of its endvertices in such a way that every two curves are either disjoint or meet only at a common endpoint. The above representation of a graph is said to be a *plane graph.*

There is a simple way of associating a topological space with a graph, which leads to another definition of planarity, trivially equivalent to the one given above. Let $p_1, p_2, \ldots$ be distinct points in $\mathbb{R}^3$, the 3-dimensional Euclidean space, such that every plane in $\mathbb{R}^3$ contains at most 3 of these points. Write (p_i, p_j) for the straight line segment with endpoints p_i and p_j (open or closed, as you like). Given a graph $G = (V, E)$, $V = (x_1, x_2, \ldots, x_n)$, the topological space

$$R(G) = \cup\{(p_i, p_j): x_i x_j \in E\} \cup \bigcup_{1}^{n} p_i \subset \mathbb{R}^3$$

is said to be a *realization* of G. A graph G is a *planar* if $R(G)$ is homeomorphic to a subset of $\mathbb{R}^2$, the plane.

Let us make some more remarks in connection with $R(G)$. A graph H is said to be a *subdivision* of a graph G or a *topological* G graph if H is obtained from G by subdividing some of the edges, that is by replacing the edges by paths having at most their endvertices in common. We shall write TG for a topological G graph. Thus TG denotes *any* member of a rather large family of graphs; for example TK^3 is an arbitrary cycle and TC^8 is an arbitrary cycle of length at least 8. It is clear that for any graph G the spaces $R(G)$ and $R(TG)$ are homeomorphic. We shall say that a graph G is *homeomorphic* to a graph H if $R(G)$ is homeomorphic to $R(H)$ or, equivalently, G and H have isomorphic subdivisions.

At the first sight one may think that in the study of planar graphs one may run into topological difficulties. This is certainly not the case. It is easily seen that the Jordan curves corresponding to the edges can be assumed to be polygons. More precisely, every plane graph is homotopic to a plane graph representing the same graph, in which the Jordan curves are piecewise linear. Indeed, given a plane graph, let $\delta > 0$ be less than half the minimal distance between two vertices. For each vertex a place a closed disc D_a of radius δ about a. Denote by J_α the curve corresponding to an edge $\alpha = ab$ and let a_α be the last point of J_α in D_a when going from a to b. Denote by J'_α the part of J_α from a_α to b_α. Let $\varepsilon > 0$ be such that if $\alpha \neq \beta$ then J'_α and J'_β are at a distance greater than 3ε. By the uniform continuity of a Jordan curve each J'_α can be approximated within ε by a polygon J''_α from a_α to b_α. To get the required piecewise linear representation of the original graph

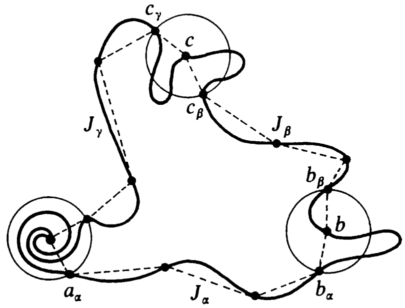


Figure I.16. Constructing a piecewise linear representation.

simply replace each J_α by the poygon obtained from J''_α by extending it in both directions by the segments aa_α and $b_\alpha b$ (see Figure I.16).

A less pedestrian argument shows that every planar graph has a *straight line representation*: it can be drawn in the plane in such a way that the edges are actually straight line segments (Exercise 28).

If we omit the vertices and edges of a plane graph G from the plane, the remainder falls into connected components, called *faces*. Clearly each plane graph has exactly one unbounded face. The *boundary* of a face is the set of edges in its closure. Since a cycle (that is a simple closed polygon) separates the points of the plane into two components, each edge of a cycle is in the boundary of two faces. A plane graph together with the set of faces it determines is called a *plane map*. The faces of a plane map are usually called *countries*. Two countries are *neighbouring* if their boundaries have an edge in common.

If we draw the graph of a convex polyhedron in the plane then the faces of the polyhedron clearly correspond to the faces of the plane graph. This leads us to another contribution of Leonhard Euler to graph theory, namely *Euler's polyhedron theorem* or simply *Euler's formula*.

Theorem 11. *If a connected plane graph G has n vertices, m edges and f faces, then*

$$n - m + f = 2.$$

Proof. Let us apply induction on the number of faces. If $f = 1$ then G does not contain a cycle so it is a tree and the result holds.

Suppose now that $f > 1$ and the result holds for smaller values of f. Let ab be an edge in a cycle of G. Since a cycle separates the plane, the edge ab is in the boundary of two faces, say S and T. Omitting ab, in the new plane graph G' the faces S and T join up to form a new face, while all other faces of G remain unchanged. Thus if n', m' and f' are the parameters of G' then $n' = n$, $m' = m - 1$ and $f' = f - 1$. Hence $n - m + f = n' - m' - f = 2$. □

Let G be a connected plane graph with n vertices, m edges and f faces; furthermore, denote by f_i the number of faces having exactly i edges in their boundaries. Of course

$$\sum_i f_i = f, \tag{3}$$

and if G has no bridge then

$$\sum_i if_i = 2m, \tag{4}$$

since every edge is in the boundary of two faces. Relations (3), (4) and Euler's formula give an upper bound for the number of edges of a planar graph of order n. This bound can be improved if the *girth* of the graph, that is the number of edges in a shortest cycle, is large. (The girth of an acyclic graph is defined to be ∞.)

Theorem 12. *A planar graph of order $n \geq 3$ has at most $3n - 6$ edges. Furthermore, a planar graph of order n and girth at least g, $3 \leq g < \infty$, has size at most*

$$\max\left\{\frac{g}{g-2}(n-2), n-1\right\}.$$

Proof. The first assertion is the case $g = 3$ of the second, so it suffices to prove the second assertion. Let G be a planar graph of order n, size m and girth at least g. If $n \leq g - 1$ then G is acyclic so $m \leq n - 1$. Assume now that $n \geq g$ and the assertion holds for smaller values of n. We may assume without loss of generality that G is connected. If ab is a bridge then $G - ab$ is the union of two vertex disjoint subgraphs, say G_1 and G_2. Putting $n_i = |G_i|$, $m_i = e(G_i)$, $i = 1, 2$, by induction we find

$$\begin{aligned} m = m_1 + m_2 + 1 &\leq \max\left\{\frac{g}{g-2}(n_1 - 2), n_1 - 1\right\} \\ &\quad + \max\left\{\frac{g}{g-2}(n_2 - 2), n_2 - 1\right\} + 1 \\ &\leq \max\left\{\frac{g}{g-2}(n-2), n-1\right\}. \end{aligned}$$

On the other hand, if G is bridgeless, (3) and (4) imply

$$2m = \sum_i if_i = \sum_{i \geq g} if_i \geq \sum_i gf_i = gf.$$

Hence, by Euler's formula,

$$m + 2 = n + f \leq n + \frac{2}{g}m,$$

and so

$$m \leq \frac{g}{g-2}(n-2). \qquad \square$$

Theorem 12 can often be used to show that certain graphs are non-planar. Thus K^5, the complete graph of order 5, is non-planar since $e(K^5) = 10 > 3(5-2)$, and $K^{3,3}$ the complete 3 by 3 bipartite graph is non-planar since its girth is 4 and $e(K^{3,3}) = 9 > (4/(4-2))(6-2)$. The non-planarity of $K^{3,3}$ implies that it is impossible to join each of 3 hours to each of 3 wells by non-crossing paths, as demanded by a well-known puzzle (see Figure 1.17).

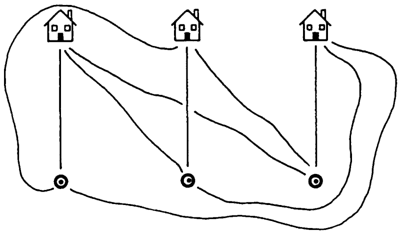


Figure I.17. Three houses and three wells.

If a graph G is non-planar then so is every topological G graph, and every graph containing a topological G graph. Thus the graphs in Figure I.18 are non-planar since they contain TK^5 and $TK^{3,3}$, respectively.

It is somewhat surprising that the converse of the trivial remarks above is also true. This beautiful result was proved by Kuratowski in 1930; as the proof is rather long though elementary, we shall not give it here.

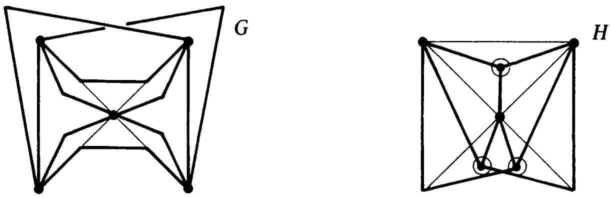


Figure I.18. G contains a TK^5 and H contains a $TK^{3,3}$.

Theorem 13. *A graph is planar iff it does not contain a subdivision of K^5 or $K^{3,3}$.* $\square$

§5 An Application of Euler Trails to Algebra

To conclude this chapter we shall show that even simple notions like the ones presented so far may be of use in proving important results. The result we are going to prove is the fundamental theorem of Amitsur and Levitzki

on polynomial identities. The *commutator* of two elements a and b of a ring S is $[a, b] = ab - ba$. Similarly, if $a_i \in S$, $1 \le i \le k$, we write

$$[a_1, a_2, \ldots, a_k] = \sum_{\sigma} (-1)^{\sigma} a_{\sigma 1} a_{\sigma 2} \ldots a_{\sigma k},$$

where the summation is over all permutations σ of the integers $1, 2, \ldots, k$. If $[a_1, a_2, \ldots, a_k] = 0$ for all $a_i \in S$, $1 \le i \le k$, then S is said to satisfy the k-th *polynomial identity*. The theorem of Amitsur and Levitzki states that the ring $M_k(R)$ of k by k matrices with entries in a commutative ring R satisfies the $2k$-th polynomial identity.

Theorem 14. *Let R be a commutative ring and let $A_1, A_2, \ldots, A_{2k} \in M_k(R)$. Then $[A_1, A_2, \ldots, A_{2k}] = 0$.*

Proof. We shall deduce the result from a lemma about Euler trails in directed multigraphs. Let $\vec{G}$ be a directed multigraph of order n with edges $e_1, e_2, \ldots, e_m$. Thus each e_i is an ordered pair of not necessarily distinct vertices. Every (directed) Euler trail P is readily identified with a permutation of $\{1, 2, \ldots, m\}$; define $\varepsilon(P)$ to be the sign of this permutation. Given not necessarily distinct vertices x, y of $\vec{G}$, put $\varepsilon(\vec{G}; x, y) = \sum_P \varepsilon(P)$, where the summation is over all Euler trails from x to y.

Lemma 15. *If $m \ge 2n$ then $\varepsilon(\vec{G}; x, y) = 0$.*

Before proving this lemma, let us see how it implies Theorem 14. Write $E_{ij} \in M_n(R)$ for the matrix whose only non-zero entry is a 1 in the ith row and jth column. Since $[A_1, A_2, \ldots, A_{2n}]$ is R-linear in each variable and $\{E_{ij}: 1 \le i, j \le n\}$ is a basis of $M_n(R)$ as an R-module, it suffices to prove Theorem 14 when $A_k = E_{i_k j_k}$ for each k. Assuming this is the case, let $\vec{G}$ be the *directed multigraph* with vertex set $\{1, 2, \ldots, n\}$, whose set of directed edges is $\{i_1 j_1, i_2 j_2, \ldots, i_{2n} j_{2n}\}$. By the definition of matrix multiplication a product $A_{\sigma 1} A_{\sigma 2} \ldots A_{\sigma 2n}$ is E_{ij} if the corresponding sequence of edges is a (directed) Euler trial from i to j and otherwise the product is 0. Hence $[A_1, A_2, \ldots, A_{2n}] = \sum_{i,j} \varepsilon(\vec{G}; i, j) E_{ij}$. By Lemma 15 each summand is 0 so the sum is also 0.

Proof of Lemma 15. We may clearly assume that $\vec{G}$ has no isolated vertices. Let $\vec{G}'$ be obtained from $\vec{G}$ by adding to it a vertex x', a path of length $m + 1 - 2n$ from x' to x and an edge from y to x' (see Figure I.19). Then

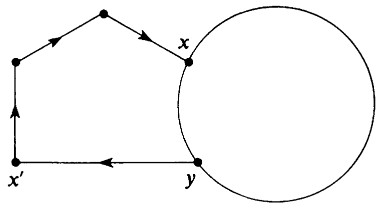


Figure I.19. The construction of $\vec{G}'$.

$\vec{G}'$ has order $n + (m + 1 - 2n) = m + 1 - n$ and size $m + m + 1 - 2n + 1 = 2(m + 1 - n)$. Furthermore, it is easily checked that $|\varepsilon(\vec{G}; x, y)| = |\varepsilon(\vec{G}'; x', x')|$. Hence it suffices to prove the theorem when $m = 2n$ and $x = y$.

Given a vertex z, let $d^+(z)$ be the number of edges starting at z and let $d^-(z)$ be the number of edges ending at z. Call $d(z) = d^+(z) + d^-(z)$ the *degree* of z and $f(z) = d^+(z) - d^-(z)$ the *flux* at z. We may assume that $\vec{G}$ contains an Euler circuit (an Euler trail from x to x), otherwise there is nothing to prove. In this case each vertex has 0 flux and even degree at least 2. Furthermore, we may assume that there is no double edge (and so no double loop) for otherwise the assertion is trivial.

In order to prove the theorem in the case $m = 2n$ and $x = y$ we apply induction on n. The case $n = 1$ being trivial, we turn to the induction step. We shall distinguish three cases.

(i) *There is a vertex $b \neq x$ of degree* 2; say $e_m = ab$ ends at b and $e_{m-1} = bc$ starts at b. If $a = c$ the assertion follows by applying the induction hypothesis to $\vec{G} - b$. If $a \neq c$ then without loss of generality $x \neq c$. Let $e_1 = cc_1$, $e_2 = cc_2, \ldots, e_t = cc_t$ be the edges starting at c. For each i, $1 \leq i \leq t$, construct a graph $\vec{G}$ from $\vec{G} - b$ by omitting e_i and adding $e_i' = ac_i$ (see Figure I.20). Then $\varepsilon(\vec{G}; x, x) = \sum_{i=1}^{t} \varepsilon(\vec{G}_i; x, x) = 0$.

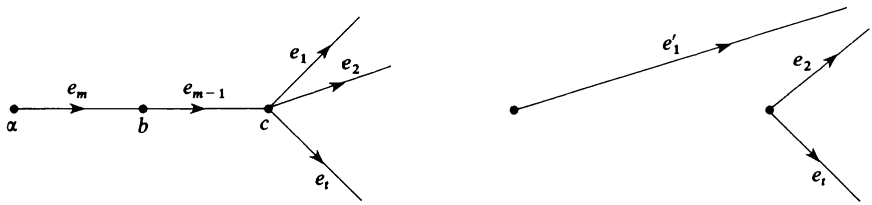


Figure I.20. The construction of $\vec{G}_1$.

(ii) *There is a loop at a vertex $b \neq x$ of degree* 4. Let e_m be the loop at b and let $e_{m-2} = ab$ and $e_{m-1} = bc$ be the other edges at b. Let $\vec{G}_0$ be obtained from $\vec{G} - b$ by adding to it an edge $e_{m-2}' = ac$. Then $\varepsilon(\vec{G}; x, x) = \varepsilon(\vec{G}_0; x, x) = 0$.

(iii) *The cases* (i) *and* (ii) *do not apply.* Since $m = 2n - \frac{1}{2}\sum_1^n d_i$ and each vertex distinct from x has degree at least 4, either each vertex has degree 4 or else $d(x) = 2$, there is a vertex of degree 6 and all other vertices have degree 4. It is easily checked (Exercise 34) that there are two adjacent vertices of degree 4, say a and b, since otherwise (ii) would hold. Now we shall apply our fourth and final graph transformation. This is more complicated than the previous ones since we shall construct two pairs of essentially different graphs from $\vec{G}$: the graphs $\vec{G}_1$, $\vec{G}_2$, $\vec{H}_6$ and $\vec{H}_7$ shown in Figure I.21. Each Euler trail from x to x in $\vec{G}$ is transformed to an Euler trail in exactly one of $\vec{G}_1$ and $\vec{G}_2$. However, the graphs $\vec{G}_2$ contain some spurious Euler trails: Euler trails that do not come from Euler trails in $\vec{G}$. As these spurious Euler

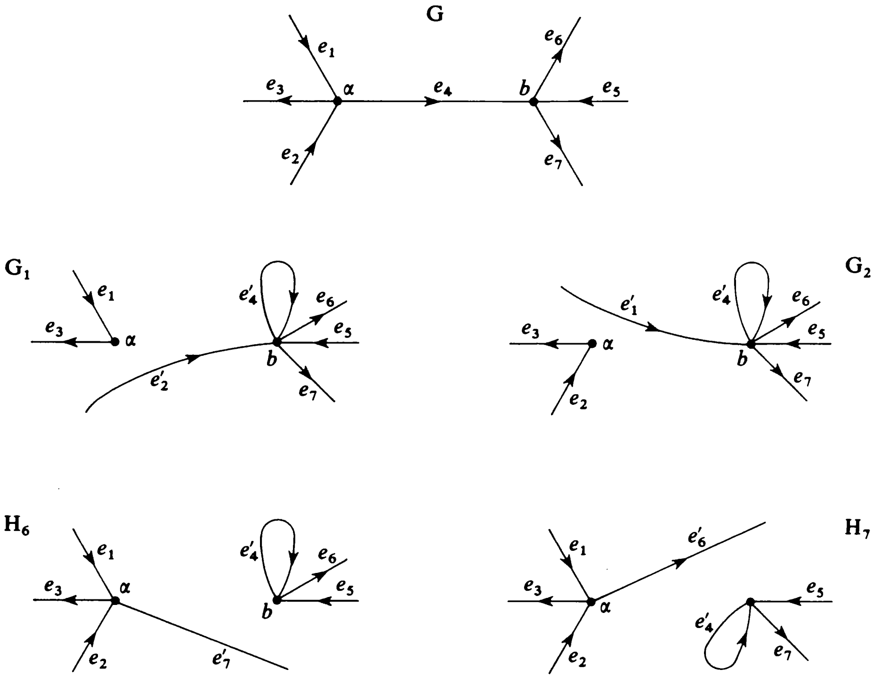


Figure I.21. The graphs $\vec{G}_1$, $\vec{G}_2$, $\vec{H}_6$ and $\vec{H}_7$.

trails are Euler trails in exactly one of $\vec{H}_6$ and $\vec{H}_7$ and they exhaust all the Euler trails of $\vec{H}_6$ and $\vec{H}_7$, we find that

$$\varepsilon(\vec{G}; x, x) = \sum_1^2 \varepsilon(\vec{G}_i; x, x) - \sum_1^2 \varepsilon(\vec{H}_i; x, x).$$

The first two terms are 0 because of (i) and the second two terms are 0 because of (ii) so $\varepsilon(\vec{G}; x, x) = 0$, completing the proof of Lemma 15. □

Exercises

1. (i) Show that every graph contains two vertices of equal degree.
   (ii) Determine all graphs with one pair of vertices of equal degree.
2.⁻ Prove that the complement of a disconnected graph is connected.
3.⁻ Show that in a graph G there exists a set of cycles such that each edge of G belongs to exactly one of these cycles iff every vertex has even degree.
4. Show that in an infinite graph G with countably many edges there exists a set of cycles such that each edge of G belongs to exactly one of these cycles iff for every $X \subset V(G)$ the set of edges joining X to $V(G) - X$ is even or infinite.
5. Show that $d_1 \le d_2 \le \cdots \le d_n$ is the degree sequence of a tree iff $d_1 \ge 1$ and $\sum_1^n d_i = 2n - 2$.

6. Show that every integer sequence $d_1 \le d_2 \le \cdots \le d_n$ with $d_1 \ge 1$ and $\sum_1^n d_i = 2n - 2k$, $k \ge 1$, is the degree sequence of a forest with k components.

7. Characterize the degree sequence of forests!

8. Prove that a regular bipartite graph of degree at least 2 does not have a bridge.

9. Let G be a graph of order n. Prove the equivalence of the following assertions.
   (i) G is a tree.
   (ii) G is connected and has $n - 1$ edges.
   (iii) G is acyclic and has $n - 1$ edges.
   (iv) $G = K^n$ for $n = 1, 2$, and if $n \ge 3$ then $G \ne K^n$ and the addition of an edge to G produces exactly one new cycle.

10. Let $\mathscr{A} = \{A_1, A_2, \ldots, A_n\}$ be a family of n (≥ 1) distinct subsets of a set X with n elements. Define a graph G with vertex set $\mathscr{A}$ in which A_iA_j is an edge iff there exists an $x \in X$ such that $A_i \triangle A_j = \{x\}$. Label the edge A_iA_j with x. For $H \subset G$ let Lab(H) be the set of labels used for edges of H. Prove that there is a forest $F \subset G$ such that Lab(F) = Lab(G).

11. (10 continued) Deduce from the result in the previous exercise that there is an element $x \in X$ such that the sets $A_1 - \{x\}, A_2 - \{x\}, \ldots, \{A_n\} - \{x\}$ are all distinct. Show that this need not hold for any n if $|\mathscr{A}| = n + 1$.

12. A *tournament* is a complete *oriented* graph, that is a directed graph in which for any two distinct vertices x and y either there is an edge from x to y or there is an edge from y to x, but not both. Prove that every tournament contains a (directed) Hamilton path.

13.⁻ Let G be a connected graph of order n and let $1 \le k \le n$. Show that G contains a connected subgraph of order k.

14.⁻ Prove that the radius and diameter of a graph satisfy the inequalities

$$\text{rad } G \le \text{diam } G \le 2 \text{ rad } G,$$

and both inequalities are best possible.

15.⁻ Given $d \ge 1$, determine

$$\max_{\text{diam } G = d} \min\{\text{diam } T\colon T \text{ is a spanning tree of } G\}.$$

16. Denote by $\beta_0(G)$ the maximum number of independent vertices in G. Prove that if G does not contain a triangle then $\Delta(G) \le \beta_0(G)$ and deduce that $e(G) \le \frac{1}{2}n\beta_0(G)$, where $n = |G|$.

17.⁻ Show that if for every vertex z of a directed graph there is an edge starting at z (that is $d^+(z) > 0$) then the graph contains a (directed) cycle.

18. A *grading* of a directed graph $\vec{G} = V, \vec{E})$ is a partitioning of V into sets $V_1, V_2, \ldots, V_k$ such that if $\overrightarrow{xy} \in \vec{E}$ then $x \in V_i$ and $y \in V_{i+1}$ for some i.

Given a directed graph G and a (non-directed) path $P = x_0x_1 \ldots x_s$, denote by $v(x_0, x_s; P)$ the number of edges $\overrightarrow{x_ix_{i+1}}$ minus the number of edges $\overrightarrow{x_{j+1}x_j}$. Prove that G has a grading iff $v(x_0, x_s; P)$ is independent of P for every pair of vertices x_0, x_s.

19. Complete the proof of Theorem 8 by showing in detail that both the second and third methods construct an economical spanning tree.

20. Show how the fourth method in Theorem 8 can be applied to find an economical spanning tree even if several edges have the same cost (cf. Figure I.4).

21. Show that every economical spanning tree can be constructed by each of the first three methods.

22. Deduce from Theorem 10 that a graph contains an Euler circuit iff all but at most one of its components are isolated vertices and each vertex has even degree. State and prove an analogous statement about the existence of an Eulerian trail from x to y.

23. The following algorithm for finding an Euler circuit $x_1 x_2 \ldots x_n$ in a graph G is due to Fleury. Pick x arbitrarily. Having chosen $x_1, x_2, \ldots, x_k$, put $G_k = G - \{x_1x_2, x_2x_3, \ldots, x_{k-1}x_k\}$. If x_k is isolated in G_k, terminate the algorithm. Otherwise let x_{k+1} be a neighbour of x_k in G_k such that $x_k x_{k+1}$ is not a bridge, unless every edge of G_k incident with x_k is a bridge.

    Prove that if G has an Euler circuit then the trail $x_1 x_2 \ldots x_l$ constructed by the algorithm is an Euler circuit.

24. A graph G is *randomly Eulerian from a vertex* x if any maximal trail starting at x is an Euler circuit. (If $T = xx_1 \ldots x_l$ then T is a maximal trail starting at x iff x_l is an isolated vertex in $G - E(T)$.) Prove that a non-empty graph G is randomly Eulerian from x iff G has an Euler circuit and x is contained in each cycle of G.

25. Let F be a forest. Add a vertex x to F and join x to each vertex of odd degree in F. Prove that the graph obtained in this way is randomly Eulerian from x.

26. Prove that a graph G is randomly Eulerian from each of two vertices x and y iff G is the union of an even number of x-y paths, any two of which have only x and y in common.

27. How would you define the number of sides of a face so that formula (4) continues to hold for graphs with bridges? Rewrite the proof of Theorem 12 accordingly.

28.⁺ Prove that every planar graph has a drawing in the plane in which every edge is a straight line segment. [*Hint*. Apply induction on the order of maximal planar graphs by omitting a suitable vertex.]

29. A plane drawing of an infinite graph is defined as that of a finite graph with the additional condition that each point has a neighbourhood containing at most one vertex and meeting only edges incident with that vertex.

    Show that Kuratowski's theorem does not hold for infinite graphs, that is construct an infinite non-planar graph without TK^5 and $TK^{3,3}$.

    Is there an infinite non-planar graph without a TK^4?

30. Let $d_1 \leq d_2 \leq \cdots \leq d_n$ be the degree sequence of a planar graph.

    (i) By making use of an upper bound for $\sum_1^n d_i$, show that if $d_1 \geq 4$ then

$$\sum_1^n d_i^2 < 2(n+3)^2 - 62.$$

(ii) Prove by induction on n that if $n \geq 4$ then

$$\sum_{1}^{n} d_i^2 \leq 2(n+3)^2 - 62.$$

Show that equality can hold for every $n \geq 4$.

31.[+] Determine the maximum of $\sum_1 d_i^2$ where $d_1 \leq d_2 \leq \cdots \leq d_n$ is the degree sequence of a planar graph with girth at least 4 (that is without triangles). What is the maximum if the girth is at least $g > 4$?

32. Fill in the small gap in the proof of Lemma 15: show that if cases (i) and (ii) do not apply then there are two adjacent vertices of degree 4.

Notes

Theorem 14 is in K. Kuratowski, Sur le problème des courbes gauches en topologie, *Fund. Math.* **15** (1930) 271–283; for a simpler proof see G. A. Dirac and S. Schuster, A theorem of Kuratowski, *Indag. Math.* **16** (1954) 343–348.

The theorem of S. A. Amitsur and J. Levitzki (Theorem 14) is in Minimal identities for algebras, *Proc. Amer. Math. Soc.* **1** (1950) 449–463; the proof given in the text is based on R. G. Swan, An application of graph theory to algebra, *Proc. Amer. Math. Soc.* **14** (1963) 367–373 and Correction to " An application of graph theory to algebra," *Proc. Amer. Math. Soc.* **21** (1969) 379–380.

CHAPTER II
Electrical Networks

This chapter is something of a diversion from the main line of the book, so some readers may wish to skip it. Only the concepts introduced in the first half of §3 will be used later, in §2 of Chapter VIII.

It does not take long to discover that an electrical network may be viewed as a graph, so the simplest problems about currents in networks are exactly questions about graphs. Does our brief acquaintance with graphs help us to tackle the problems? As it will transpire in the first section, the answer is yes; for after a short review of the basic ideas of electricity we make use of spanning trees to obtain solutions. Some of these results can be reformulated in terms of tilings of rectangles and squares, as we shall show in §2. The last section introduces elementary algebraic graph theory which is then applied to electrical networks.

It should be emphasised that in the problems we consider we use hardly more than the terminology of graph theory; virtually the only concept to be used is that of a spanning tree.

§1 Graphs and Electrical Networks

A simple electrical network can be regarded as a graph in which each edge e_i has been assigned a real number r_i called its *resistance*. If there is a *potential difference* p_i between the endvertices of e_i, say a and b, then an *electrical current* w_i will flow in the edge e_i from a to b according to *Ohm's law*:

$$w_i = \frac{p_i}{r_i}.$$

Though to start with we could restrict our attention to electrical networks corresponding to graphs, in the simplifications that follow it will be essential to allow *multiple* edges, that is to consider *multigraphs* instead of graphs. Furthermore, we *orient* each edge arbitrarily from one endvertex to the other so that we may use p_i to denote the *potential difference in the edge* e_i, meaning the difference between the potentials of the *initial vertex* and the *endvertex*. Similarly w_i is the *current in the edge* e_i, meaning the current in e_i in the direction of the edge. (Note that we regard a negative current $-w_i$ as a positive current w_i in the other direction.) Thus, throughout the section we consider *directed multigraphs*, that is directed graphs which may contain *several* edges directed from a to b. However, in this section there is no danger of confusion if we use ab to denote an edge *from* a *to* b; in the next section we shall be more pedantic. Thus

$$w_{ab} = -w_{ba} \quad \text{and} \quad p_{ab} = -p_{ba}.$$

In many practical problems electrical currents are made to enter the network at some points and leave it at others, and we are interested in the consequent currents and potential differences in the edges. These are governed by the famous laws of Kirchhoff (another renowned citizen of Königsberg).

Kirchhoff's potential (or *voltage*) *law* states that the potential differences round any cycle $x_1x_2 \dots x_k$ sum to 0:

$$p_{x_1x_2} + p_{x_2x_3} + \cdots + p_{x_{k-1}x_k} + p_{x_kx_1} = 0.$$

Kirchhoff's current law postulates that the total current outflow from any point is 0:

$$w_{ab} + w_{ac} + \cdots + w_{au} + w_{a\infty} = 0.$$

Here $ab, ac, \dots, au$ are the edges incident with a, and $w_{a\infty}$ denotes the amount of current that leaves the network at a. (In keeping with our convention, $w_{\infty a} = -w_{a\infty}$ is the amount of current entering the network at a). For vertices not connected to external points we have

$$w_{ab} + w_{ac} + \cdots + w_{au} = 0.$$

Note that if we know the resistances, then the potential law can be rewritten as a restriction on the currents in the edges. Thus we may consider that the currents are governed by the Kirchhoff laws only; the physical characteristics of the network (the resistances) affect only the parameters in these laws.

It is also easily seen that the potential law is equivalent to saying that one can assign absolute potentials $V_a, V_b, \dots$ to the vertices $a, b, \dots$ so that the potential difference between a and b is $V_a - V_b = p_{ab}$. If the network is connected and the potential differences p_{ab} are given for the edges, then we are free to choose arbitrarily the potential of one of the vertices, say V_a, but then all the other potentials are determined. In this section we shall work

with the potentials, usually choosing the potential of one of the vertices to be 0, but we must keep in mind that this is the same as the application of the voltage law.

In the most fundamental problems current is only allowed to enter the network at a single point s, the *source*, and only leave it at another point t, the *sink*. (We shall indicate later that the general problem can be reduced to these fundamental problems.) If the size of the current from s to t is w and the potential difference between s and t is p, then by Ohm's law $r = p/w$ is the *total resistance* of the network between s and t. As an example of the use of the Kirchhoff laws we shall evaluate the total resistance between s and t of the simple network shown in Figure II.1.

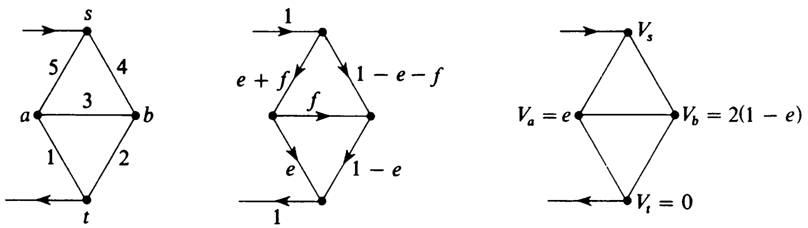


Figure II.1. The resistances, the currents and the potentials.

This network has 5 resistors, of values 1, 2, 3, 4 and 5 ohms, as shown in the first picture. If we suppose that a unit current flows into the system at s and leaves it at t, then the consequent edge currents must be as in the second picture, for suitable values of e and f. Finally, the potentials $V_t = 0$, V_a, V_b, V_s assigned to the vertices must satisfy Ohm's law, so $V_a = 1 \cdot e = e$, $V_b = 2(1 - e)$ and $V_s = V_a + 5(e + f) = 6e + 5f$. Ohm's law has to be satisfied in two more edges, ab and bs, giving us

$$V_a = e = V_b + 3f = 2(1 - e) + 3f$$

and

$$V_s = 6e + 5f = V_b + 4(1 - e - f) = 2(1 - e) + 4(1 - e - f).$$

Hence

$$e = 2 - 2e + 3f$$

and

$$6e + 5f = 6 - 6e - 4f,$$

giving $e = 4/7$, $f = -2/21$ and $V_s = 6e + 5f = 62/21$. In particular, the total resistance from s to t is $(V_s - V_t)/w = 62/21$.

The calculations are often simplified if we note that Kirchhoff's equations are linear and homogeneous in all currents and potential differences. This implies the so-called *principle of superposition*: any combination of solutions is again a solution. As an application of the principle of superposition one

can show that any current resulting from multiple sources and sinks can be obtained by superposing flows belonging to one source and one sink, that is solutions of the fundamental problems mentioned above can be used to solve the general problem. Furthermore, the principle of superposition implies immediately that there is at *most one solution*, no matter how the sources and sinks are distributed. Indeed, the difference of two distinct solutions is a flow in which *no* current enters or leaves the network at any point. If in this flow there is a positive current in some edge from a to b, then by the current law a positive current must go from b to c, then from c to d, etc., giving a trail $abcd\dots$. Since the network is finite, this trail has to return to a point previously visited. Thus we obtain a circuit in whose edges positive currents flow in one direction. But this is impossible since it implies that the potential of each vertex is strictly greater than that of the next one round the circuit.

Before proving the *existence* of a solution (which is obvious if we believe in the physical interpretation), we shall calculate the total resistance of two networks. Unless the networks are very small, the calculations can get very heavy, and electrical engineers have a number of standard tricks to make them easier.

The very simple networks of Figure II.2 show two resistors r_1 and r_2 connected first *in series* and then *in parallel*. Let us put a current of size 1

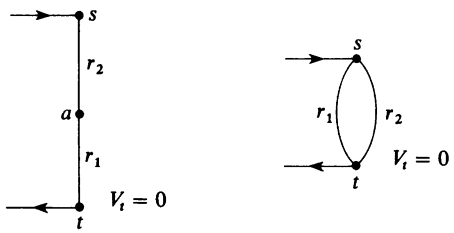


Figure II.2. Resistors connected in series and in parallel.

through the networks, from s to t. What are the total resistances? In the first case

$$V_a = r_1 \quad \text{and} \quad V_s = V_a + r_2 = r_1 + r_2$$

so the total resistance is

$$r = r_1 + r_2.$$

In the second case, when they are connected in parallel, if a current of size e goes through the first resistor and so a current of size $1 - e$ through the second, then

$$V_s = r_1 e = r_2(1 - e), \quad \text{so} \quad e = \frac{r_2}{r_1 + r_2}$$

and the total resistance is given by

$$r = \frac{r_1 r_2}{r_1 + r_2} \quad \text{or} \quad \frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2}.$$

This indicates that reciprocals of resistances, or *conductances*, are just as natural as the resistances themselves, and indeed are slightly more convenient in our presentation, especially in §3. (The conductance of an edge of resistance 1 ohm is 1 mho). What we have shown now is that for series connection the *resistances add* and for parallel connection the *conductances add*.

The use of conductances is particularly convenient when considering certain limiting cases of Ohm's law. If the resistance of an edge ab is 0, then we necessarily have $V_a = V_b$, and from an electrical point of view the vertices can be regarded as identical. In the usual slang, a has been "shorted" (short-circuited) to b. Of course, a may be shorted to b if there is some other reason why $V_a = V_b$. At the other extreme, we can introduce edges of 0 conductance without effecting the currents and potentials. Conversely, we make an edge have 0 conductance by "cutting" it. Of course, an edge of 0 resistance is said to have ∞ conductance and an edge of 0 conductance is said to have ∞ resistance.

Let us see now how the acquaintance with resistors in series and in parallel and the possibility of shorting vertices can help us to determine the total resistance. As an example, let us take the network formed by the edges of a cube, in which each edge has 1 ohm resistance. What is the total resistance across an edge st? Using the notation of the first picture in Figure II.3, we see that by symmetry $V_a = V_c$ and $V_d = V_f$, so c can be shorted to a and f to d, giving us the second picture. From now on we can simplify resistors connected parallel and in series, until we find that the total resistance is $\frac{7}{12}$. Knowing this, it is easy to recover the entire current flow.

Another important device in practical calculations is the so-called *star-delta transformation*. If a vertex v is joined to just three vertices, say

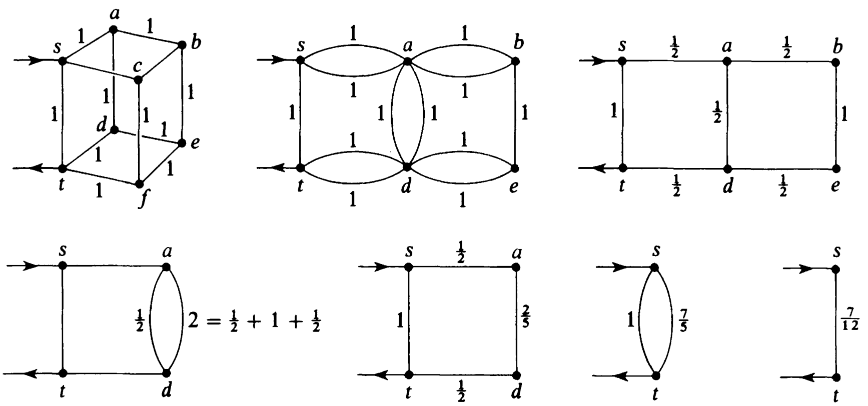


Figure II.3. Calculating the total resistance of a cube.

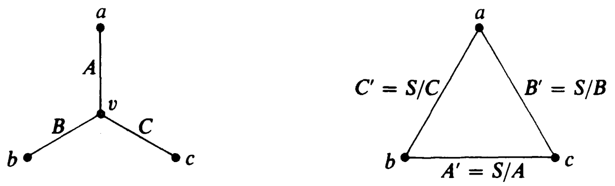


Figure II.4. The star-delta transformation; $S = AB + BC + CA$.

a, b and c, by edges of resistances A, B and C, then we call v the center of a *star*, as in the first picture of Figure II.4. If no current is allowed to enter or leave at v, then we are allowed to replace this star by the *delta*-configuration shown in the second picture of Figure II.3, because, as the reader should check, these two networks have the same total resistances across all three pairs ab, bc and ca. Of course we may apply the transformation in reverse, replacing A', B', C' by $A = B'C'/T$, $B = C'A'/T$ and $C = A'B'/T$, where $T = A' + B' + C'$. Incidentally, the formulae become symmetrical if we use resistances in the first transformation and conductances in the second: $A' = B + C + BC/A$, and $\alpha = \beta' + \gamma' + \beta'\gamma'/\alpha'$, where $\alpha, \beta, \ldots$ are the conductances.

As an application of the star-delta transformation, let us calculate the total resistance of a tetrahedron across an edge, in which the resistances are as in Figure II.5. The pictures speak for themselves.

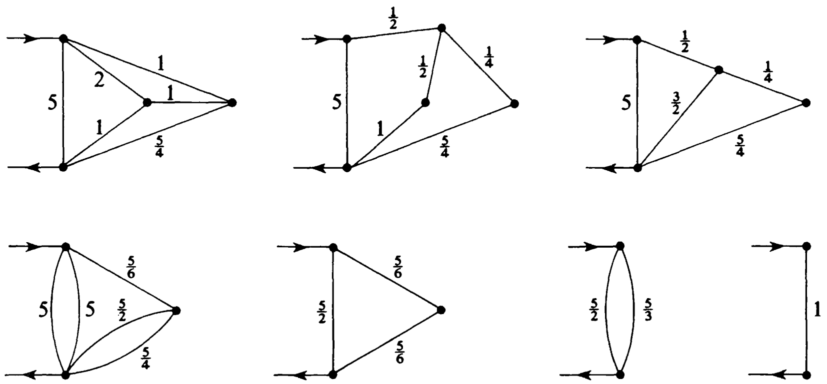


Figure II.5. Applications of the star-delta transformation.

We shall conclude this section on a slightly more theoretical note: we shall prove the existence of a solution. More precisely, we shall show that if a current of size 1 is put through a network then the current in an edge can be expressed in terms of the numbers of certain spanning trees. For simplicity we assume that the graph G of the network is connected, each edge has unit resistance, and a current of size 1 enters at a vertex s and leaves at t.

Theorem 1. *Given an edge ab, denote by $N(s, a, b, t)$ the number of spanning trees of G in which the (unique) path from s to t contains a and b, in this order. Define $N(s, b, a, t)$ analogously and write N for the total number of spanning trees. Finally, let $w_{ab} = \{N(s, a, b, t) - N(s, b, a, t)\}/N$.*

Distribute currents in the edges of G by sending a current of size w_{ab} from a to b for every edge ab. Then there is a total current size 1 from s to t satisfying the Kirchhoff laws.

PROOF. For each spanning tree T there is exactly one neighbour x_T of s that is on the s-t path P_T contained in T. Hence $\sum_{b \in \Gamma(s)} N(s, s, b, t) = N$, where $\Gamma(s)$ is the set of neighbours of s. Since $N(s, b, s, t) = 0$ for every vertex $b \in \Gamma(s)$, we find that $\sum_{b \in \Gamma(s)} w_{sb} = 1$. By symmetry $\sum_{a \in \Gamma(t)} w_{at} = 1$; so Kirchhoff's current law is satisfied at s and t provided a current of size 1 enters the network at s and leaves it at t.

To prove the theorem we have to show that if no current is allowed to enter or leave the network at any other point then the current and potential laws are satisfied.

Let us take the current law first; to simplify the situation multiply all currents by N. What is the contribution of a spanning tree T to the current entering and leaving a vertex y distinct from s and t? If y is not on the s-t path P_T then T contributes nothing to $N(s, x, y, t)$ or $N(s, y, x, t)$. Now if y is on the s-t path P_T, say $P_T = s \ldots xyz \ldots t$, then T contributes 1 to $N(s, y, z, t)$ and it also contributes 1 to $N(s, x, y, t)$. In other words T contributes a current of size 1 *from* y and a current of size 1 *into* y. Thus the current law is satisfied.

As all edges have the same resistance, the potential law claims that the total current in a cycle with some orientation is zero. To show this we proceed as earlier, but first reformulate slightly the definition of $N(s, a, b, t)$. Call a forest F a *thicket* if it has exactly two components, say F_s and F_t, such that s is in F_s and t is in F_t. Then $N(s, a, b, t)$ is the number of thickets $F = F_s \cup F_t$ for which $a \in F_s$ and $b \in F_t$ and $N(s, b, a, t)$ is defined analogously. What is then the contribution of a thicket $F = F_s \cup F_t$ to the total current in a cycle? It is the number of cycle edges from F_s to F_t minus the number of cycle edges from F_t to F_s; so it is zero. □

The proof above can be rewritten word for word to give a solution in the case when the edges have arbitrary conductances. For a spanning tree T define the *weight* $w(T)$ of T as the product of the conductances of its edges. Let N^* be the sum of the weights of all the spanning trees, let $N^*(s, a, b, t)$ be the sum of the weights of all the spanning trees in which b follows a on the (unique) s-t path in the tree, and let $N^*(s, b, a, t) = N^*(t, a, b, s)$.

Theorem 2. *There is a distribution of currents satisfying Kirchhoff's laws in which a current of size 1 enters at s and leaves at t. The value of the current in and edge ab is given by $\{N^*(s, a, b, t) - N^*(s, b, a, t)\}/N.^*$* □

Corollary 3. *If the conductances of the edges are rational and a current of size 1 goes through the network then the current in each edge has rational value.* □

§2 Squaring the Square

This is a diversion within a diversion; we feel bound to draw attention to a famous problem arising from recreational mathematics which is related to the theory of electrical networks. Is there a *perfect squared square*? In other words, is it possible to subdivide a closed square into finitely many (but at least two) square regions of distinct sizes that intersect only at their boundaries?

Let us consider a squared rectangle, say the one shown in Figure II.6; the number in a square is the length of its side. Let us cut this rectangle out of a sheet of nichrome (or any other material with low conductivity) and let us put rods made of silver (or some other material of high conductivity) at the top and bottom.

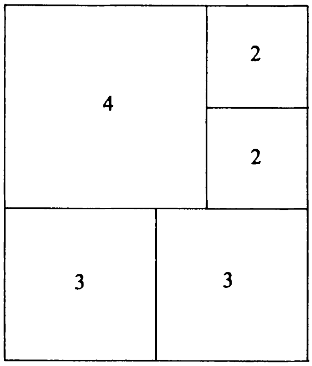


Figure II.6. A (non-perfect) squaring of a 6×7 rectangle.

What happens if we ensure that the silver rod at the top is at 7 volts while the rod at the bottom is kept at 0? Of course a uniform current will flow from top to bottom. In fact the potential at a point of the rectangle will depend only on the height of the point: the potential at height x will be x volts. Furthermore, there will be no current across the rectangle, only from top to bottom. Thus the current will not change at all if (i) we place silver rods on the horizontal sides of the squares and (ii) cut narrow slits along the vertical sides, as shown in the first picture of Figure II.7.

Now since silver is a very good conductor, the points of each silver rod have been shortened, so can be identified. Thus as an electric conductor the whole rectangle behaves like the *plane* network shown in the second picture of Figure II.7, in which the conductance of an edge is equal to the conductance

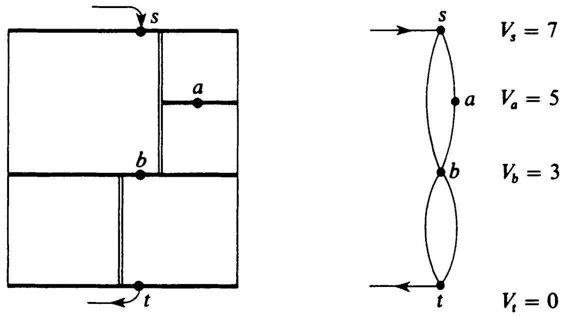


Figure II.7. The electrical network associated with our rectangle.

of the corresponding square from top to bottom. Clearly the conductance of a rectangle from top to bottom is proportional to the length of a horizontal side and the resistance is proportional to a vertical side. Consequently all squares have the same resistance, say unit resistance, so all edges in Figure II.7 have unit resistance. What is the potential drop in an edge? It is the side length of the corresponding square. What is the resistance of the whole system? The ratio of the horizontal side of the original big rectangle to the vertical side, that is $\frac{7}{6}$.

Since the process above is reversible, that is every squared rectangle can be obtained from some network we have an effective *tool* to help us in our search for squared squares. Take a connected planar graph G and turn it into an electrical network by giving each edge resistance 1. Calculate the total resistance from a vertex s to a vertex t. If this is also 1, the network may correspond to a suitably squared square. If the potential differences in the edges are all distinct, all squares have different sizes, so we have a perfect squared square.

Of course, at this stage our problem is far from being solved; we do not even know that there *must* exist a squared square. However, we have a chance to search systematically for a solution. Many squared squares have been

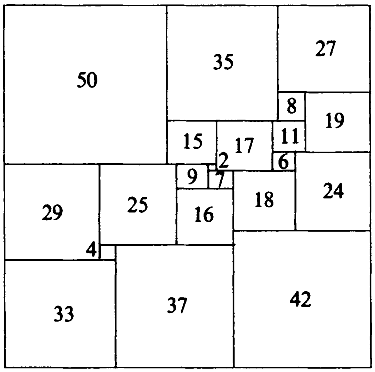


Figure II.8. A perfect squared square: a tiling of a square with 21 incongruent squares.

found with the help of computers, but the first examples were found without computers by Sprague in 1939 and by Brooks, Smith, Stone and Tutte in 1940. The smallest number of squares that can tile a square is 21; Figure II.8 shows this tiling, due to Duijvestijn. In fact, this is the only tiling of order 21.

The connection between squaring a rectangle and electrical networks gives us immediately a beautiful result first proved by Dehn in 1903. Corollary 3 tells us that if each edge has resistance 1 and a current of size 1 flows through the system then in each edge the value of the current is rational. This translates to the following result about squared rectangles.

Theorem 4. *If a rectangle can be tiled with squares then the ratio of two neighbouring sides of the rectangle is rational.* □

Equivalently: a rectangle can be tiled with squares iff it can be tiled with congruent squares.

It is easily seen that electrical networks can be used to obtain tilings of rectangles of prescribed shapes: an edge e of resistance r corresponds to a rectangle in which the height is r times the base. The reader is encouraged to solve the exercises based on this simple idea (Exercises 9–12); some other exercises concern related packing problems (Exercises 13–15).

§3 Vector Spaces and Matrices Associated with Graphs

The vertex space $C_0(G)$ of a graph G is the complex vector space of all functions from $V(G)$ into $\mathbb{C}$. Similarly the *edge space* $C_1(G)$ is the complex vector space of all functions from $E(G)$ into $\mathbb{C}$. In these definitions it is sometimes convenient to replace the complex field by F_2, the field of order 2, or by other fields. We shall take $V(G) = \{v_1, v_2, \ldots, v_n\}$ and $E(G) = \{e_1, e_2, \ldots, e_m\}$ so that dim $C_0(G) = n$ and $C_1(G) = m$. The elements of $C_0(G)$ are usually written in the form $\mathbf{x} = \sum_{i=1}^n x_i v_i$ or $\mathbf{x} = (x_i)_1^n$. The sum $\sum_{i=1}^n x_i v_i$ is a *formal sum* of the vertices but if we think of v_i as the function $V(G) \to \mathbb{C}$ which is 0 everywhere, except at the vertex v_i, where it is 1, then $v_1, \ldots, v_n$ is a basis of $C_0(G)$ and the sum above simply expresses an element in terms of the basis elements. Similarly an element of $C_1(G)$ may be written as $\mathbf{y} = \sum_{i=1}^m y_i e_i$ or $\mathbf{y} = (y_i)_1^m$. We call $(v_1, \ldots, v_m)$ the *standard basis* of the vertex space $C_0(G)$ and $(e_1, \ldots, e_m)$ the *standard basis* of the edge space. We shall endow these spaces with the inner product in which the standard bases are orthonormal: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_i x_i \bar{y}_i$.

In this section we shall be concerned mostly with the edge space $C_1(G)$; to start with we define two subspaces which will turn out to be orthogonal complements of each other. Let L be a cycle in G with a given cyclic orientation $L = u_1 u_2 \ldots u_l$. If $e_i = u_j u_{j+1}$ and e_i is oriented from u_j to u_{j+1} then

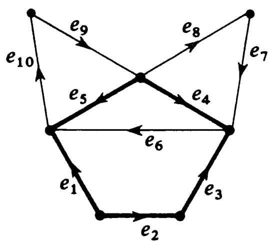


Figure II.9. If the thick cycle L is oriented anti-clockwise, its vector in $C_1(L)$ is $\mathbf{z}_L = (-1, 1, 1, -1, 1, 0, \ldots, 0)$.

we say that e_i *is oriented as* L. This oriented cycle L can be identified with an element $\mathbf{z}_L$ of $C_1(G)$:

$$\mathbf{z}_L(e_i) = \begin{cases} 1 & \text{if } e_i \in E(L) \text{ and } e_i \text{ is oriented as } L, \\ -1 & \text{if } e_i \in E(L) \text{ and } e_i \text{ is not oriented as } L, \\ 0 & \text{if } e_i \notin E(L). \end{cases}$$

A simple example is shown in Figure II.9. Denote by $Z(G)$ the *subspace* of $C_1(G)$ spanned by the vectors $\mathbf{z}_L$ as L runs over the set of cycles; $Z(G)$ is the *cycle space* of G.

Let now P be a partition $V = V_1 \cup V_2$ of the vertex set of G. Consider the set (V_1, V_2) of edges from V_1 to V_2; such a set of edges is called a *cut*. There is a vector $\mathbf{u}_p$ in $C_1(G)$ called a *cut* vector, naturally associated with this partition P:

$$\mathbf{u}_p(e_i) = \begin{cases} 1 & \text{if } e_i \text{ goes from } V_1 \text{ to } V_2 \\ -1 & \text{if } e_i \text{ goes from } V_2 \text{ to } V_1 \\ 0 & \text{if } e \notin E(V_1, V_2). \end{cases}$$

We write $U(G)$ for the *subspace* of the edge space $C_1(G)$ spanned by all the cut vectors $\mathbf{u}_p$; $U(G)$ is the *cut* (or *cocycle*) space of G.

Theorem 5. *The inner product space $C_1(G)$ is the orthogonal direct sum of the cycle space $Z(G)$ and the cut space $U(G)$. If G has n vertices, m edges and k components then*

$$\dim Z(G) = m - n + k \quad \textit{and} \quad \dim U(G) = n - k.$$

Proof. Let us see first that $Z(G)$ and $U(G)$ are orthogonal. Let L be a cycle and P a partition $V = V_1 \cup V_2$. What is the product $\langle \mathbf{z}_L, \mathbf{u}_p \rangle$? It is simply the number of edges of L going from V_1 to V_2 in the orientation of L, minus the number of edges of L from V_2 to V_1. Thus $\langle \mathbf{z}_L, \mathbf{u}_p \rangle = 0$ for every cycle L and partition P, so $Z(G)$ and $U(G)$ are indeed orthogonal.

Since the dimension of $C_1(G)$ is the number of edges, m, both assertions will be proved if we show that $\dim Z(G) \geq m - n + 1$ and $\dim U(G) \geq n - 1$. We shall first prove this under the assumption that G is connected; the general case will follow easily.

Thus let us assume that G is connected, that is $k = 1$. Let T be a spanning tree of G. We shall make use of T to exhibit $m - n + 1$ independent vectors in $Z(G)$ and $n - 1$ independent vectors in $U(G)$. We may choose the indices of the edges in such a way that $e_1, e_2, \ldots, e_{n-1}$ are the tree edges and $e_n, e_{n+1}, \ldots, e_m$ are the remaining edges, the *chords* of T.

We know that for every chord e_i $(i \geq n)$ there is a (unique) oriented cycle C_i such that $\mathbf{z}_{C_i}(e_i) = 1$ and $\mathbf{z}_{C_i}(e_j) = 0$ for every other chord e_j, that is whenever $j \geq n$ and $j \neq i$. (Shortly: $\mathbf{z}_{C_i}(e_j) = \delta_{ij}$ if $j \geq n$, where δ_{ij} is the Kronecker delta.) We call C_i the *fundamental cycle* belonging to e_i (with respect to T); $\mathbf{z}_{C_i}$ is a *fundamental cycle vector* (see Figure II.10). Similarly by

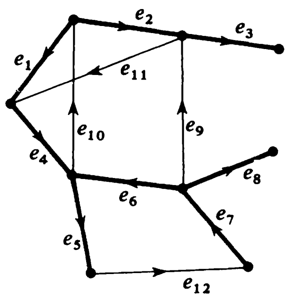


Figure II.10. The fundamental cycle vector belonging to e_9 is $\mathbf{z}_{C_i} = e_9 - e_2 + e_1 + e_4 - e_6$, the fundamental cut vector belonging to e_4 is $\mathbf{u}_{p_4} = e_4 - e_{10} - e_9$.

deleting an edge e_i of T the remainder of the spanning tree falls into two components. Let V_1^i be the vertex set of the component containing the initial vertex of e_i and let V_2^i be the vertex set of the component containing the terminal vertex of e_i. If p_i is the partition $V = V_1^i \cup V_2^i$ then clearly $\mathbf{u}_{p_i}(e_j) = \delta_{ij}$ for $1 \leq j \leq n - 1$. The cut $E(V_1^i, V_2^i)$ is the *fundamental* cut belonging to e_i (with respect to T) and $\mathbf{u}_{p_i}$ is the *fundamental cut vector*.

It is easily seen that $\{\mathbf{z}_{C_i}: n \leq i \leq m\}$ is an independent set of cycle vectors. Indeed, if $\mathbf{z} = \sum_{i=n}^{m} \lambda_i \mathbf{z}_{C_i} = 0$ then for every $j \geq n$ we have $0 = \mathbf{z}(e_j) = \sum_{i=n}^{m} \lambda_i \delta_{ij} = \mathbf{y}_j$ and so every coefficient λ_j is 0. Similarly the fundamental cut vectors $\mathbf{u}_{p_i}$, $1 \leq i \leq n - 1$, are also independent. Hence $\dim Z(G) \geq m - n + 1$ and $\dim U(G) \geq n - 1$, as required.

Finally, the general case $k \geq 1$ follows immediately from the case $k = 1$. For if G has components $G_1, G_2, \ldots, G_k$ then $C_1(G)$ is the orthogonal direct sum of the subspaces $C_1(G_i)$, $i = 1, 2, \ldots, k$; furthermore $Z(G_i) = Z(G) \cap C_1(G_i)$ and $U(G_i) = U(G) \cap C_1(G_i)$. □

The proof above shows that $\dim Z(G)$, called the *cyclomatic* number of G, and $\dim U(G)$ are independent of the field over which the edge space is defined. The use of a spanning tree in the proof is not compulsory; in some cases, for instance in the case of a planar graph, there are other natural cycle and cut bases (cf. Exercise 16).

There are several matrices naturally associated with a graph and its vector spaces discussed above. The *adjacency matrix* $A = A(G) = (a_{ij})$ of a graph G is the $n \times n$ matrix given by

$$a_{ij} = \begin{cases} 1 & \text{if } v_i v_j \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

In order to define the incidence matrix of a graph we again consider an *orientation* of the edges, as in the definition of the cycle and cut spaces. The *incidence matrix* $B = B(G) = (b_{ij})$ of G is the $n \times m$ matrix defined by

$$b_{ij} = \begin{cases} 1 & \text{if } v_i \text{ is the initial vertex of the edge } e_j, \\ -1 & \text{if } v_i \text{ is the terminal vertex of the edge } e_j, \\ 0 & \text{otherwise.} \end{cases}$$

There is a simple connection between the two matrices A and B.

Theorem 6. *$BB^t = D - A$, where B^t denotes the transpose of B and D is the $n \times n$ diagonal matrix in which $(D)_{ii}$ is $d(v_i)$, the degree of the vertex v_i in G.*

PROOF. What is $(BB^t)_{ij}$? It is $\sum_{l=1}^m b_{il} b_{jl}$, which is $d(v_i)$ if $i = j$, -1 if $v_i v_j$ is an edge (if $e_l = v_i v_j$ is directed from v_i to v_j, then $b_{il} b_{jl} = 1(-1) = -1$ and all other products are 0), and 0 if $v_i v_j$ is not an edge and $i \neq j$. □

We may and will identify the matrices A and B with the linear maps $A: C_0(G) \to C_0(G)$ and $B: C_1(G) \to C_0(G)$ they define in the standard bases: $(A\mathbf{x})_i = \sum_{j=1}^n a_{ij} x_j$ and $(B\mathbf{y})_i = \sum_{j=1}^m b_{ij} y_j$. If we wanted to be pedantic, we would write the vectors in the vertex and edge spaces as *column vectors*, or we would put $A\mathbf{x}^t$ and $B\mathbf{y}^t$, where t stands for transposition; we shall not do this since there is no danger of confusion. If C is a cycle then clearly $B\mathbf{z}_C = 0 \in C_0(G)$; in fact, it is easily shown (cf. Exercise 17) that the cycle space is exactly the *kernel* of B. Thus the rank of B is $r(B) = m - (m - n + k) = n - k$. Furthermore, the transpose of B maps $C_0(G)$ into $C_1(G)$, and the *image* of B^t is exactly the cut space (cf. Exercise 18).

In Chapter VIII we shall discuss in some detail the eigenvalues and eigenvectors of the adjacency matrix; in this section we shall use the matrices to solve the electrical network problem discussed in the first section. In fact, it was Kirchhoff who first realized the applicability of matrix algebra to graph theory, exactly in connection with the electrical network problem.

How can we formulate the Kirchhoff laws in terms of matrices and vectors in the edge space? Let us assume that G' is the graph of our electrical network; $V(G') = \{v_1, v_2, \ldots, v_{n-1}\}$, $E(G') = \{e_1, e_2, \ldots, e_{m'}\}$, the network is connected and we have a voltage generator ensuring that the potential difference between v_i and v_j is $g_i - g_j$ volts for $1 \leq i < j \leq k$. In order to express Kirchhoff's laws in a neat form, we add a vertex v_n to G', and join it to $v_1, v_2, \ldots, v_k$; the new graph is G. Let $m = m' + k$ and $e_{m'+1} = v_n v_i$, $i = 1, 2, \ldots, k$, so that $V(G) = \{v_1, c_2, \ldots, v_n\}$ and $E(G) = \{e_1, e_2, \ldots, e_m\}$.

Give the edges of G' an arbitrary orientation and let w_i be the amount of current flowing in the edge e_i in the direction of e_i; thus $w_i = -1$ means a current of 1 ampere in the opposite direction. Direct each new edge $e_{m'+i}$ from v_n to v_i and let $w_{m'+i}$ be the *total current* entering the network at v_i. Once again, $w_{m'+i} = -1$ means that a current of 1 ampere leaves the network at v_i. The vector $\mathbf{w} = (w_1, w_2, \ldots, w_m) \in C_1(G)$ is the *current vector*. In this notation Kirchhoff's current law takes the form

$$B\mathbf{w} = \mathbf{0}. \tag{1}$$

We have no greater difficulty in formulating Kirchhoff's potential law in matrix form. Let p_i be the potential difference in the edge e_i and let $\mathbf{p} = (p_1, p_2, \ldots, p_m) \in C_1(G)$ be the *potential vector*. The potential law states that $\langle \mathbf{z}, \mathbf{p} \rangle = 0$ for every cycle $\mathbf{z} \in C_1(G)$. Instead of postulating this about every cycle, we collect all necessary information into a single matrix. As before, we choose a spanning tree T in G and label the edges so that $e_1, e_2, \ldots, e_{n-1}$ are the tree edges and $e_n, e_{n+1}, \ldots, e_m$ are the chords. Let C be the $m \times (m - n + 1)$ matrix whose ith column is the fundamental cycle vector $\mathbf{z}_{C_{n-1+i}}$ belonging to the edge e_{n-1+i}, $i = 1, 2, \ldots, m - n + 1$. Since the fundamental cycle vectors form a basis of the cycle space, the potential law takes the form

$$C^t\mathbf{p} = \mathbf{0}, \tag{2}$$

where C^t denotes the transpose of C.

Now in order to find the current through the edges of G' we need one more equation, namely the equation relating the potential to the current, the conductance and the voltage generator. For $i \leq m'$ let d_i be the conductance of the edge e_i and postulate that each new edge e_j ($j \geq m' + 1$) has conductance $d_j = 0$. We may assume that $d_i > 0$ for every $i \leq m'$ since otherwise the edge e_i could have been cut. Let D be the $m \times m$ diagonal matrix with $(D)_{ii} = d_i$. Finally, let $\mathbf{g} = (0, \ldots, 0, g_1, g_2, \ldots, g_k) \in C_1(G)$ be the vector of the voltage generator. Then clearly

$$\mathbf{p} = D\mathbf{w} + \mathbf{g}. \tag{3}$$

This equation contains all the information we have about the electric current in addition to the Kirchhoff laws.

In order to solve (1), (2) and (3) for $\mathbf{w}$ and $\mathbf{p}$, we shall split $C_1(G)$ as $E_T + E_N$, where E_T is the subspace spanned by the tree edges and E_N is spanned by the chords, the edges not belonging to T. Let $\mathbf{w} = (\mathbf{w}_T, \mathbf{w}_N)$ and $\mathbf{p} = (\mathbf{p}_T, \mathbf{p}_N)$ be the corresponding splittings; furthermore, writing $\tilde{B}$ for the matrix obtained from B by omitting the last row, we have

$$C = \begin{pmatrix} C_T \\ C_N \end{pmatrix} \quad \text{and} \quad \tilde{B} = (B_T B_N).$$

As the columns of C are the fundamental cycles, C_N is the $(m - n + 1) \times (m - n + 1)$ identity matrix I_{m-n+1}. Since the kernel of B contains the cycles,

$BC = 0$ and so $\tilde{B}C = 0$, giving $B_T C_T = -B_N$. Now B_T is invertible, as the reader should check (cf. Exercise 19), so

$$C_T = -B_T^{-1} B_N.$$

After this preparation we can easily solve our equations.

Theorem 7. *The electric current* $\mathbf{w}$ *satisfying* $p = D\mathbf{w} + g$ *is given by* $\mathbf{w} = -C(C^t DC)^{-1} C^t g$.

PROOF. Equation (1) implies that $B_T \mathbf{w}_T + B_N \mathbf{w}_N = \mathbf{0}$, so $\mathbf{w}_T = -B_T^{-1} B_N \mathbf{w}_N = C_T \mathbf{w}_N$. Hence $\mathbf{w} = C\mathbf{w}_N$. Combining (2) and (3) we find that $C^t D\mathbf{w} + C^t \mathbf{g} = \mathbf{0}$ and so $(C^t DC)\mathbf{w}_N = -C^t \mathbf{g}$. As $C^t DC$ is easily shown to be invertible, the result follows. □

Clearly Theorem 7 is valid in a somewhat more general situation, not only when G and $\mathbf{g}$ are defined as above. In fact, the following conditions are sufficient (and more or less necessary) for the existence of a unique current $\mathbf{w}$: $g_i d_i = 0$ for every i and the edges e_j with $d_j > 0$ form a connected subgraph.

Furthermore, the results hold for multigraphs, that is "graphs" in which several edges may join the same pair of vertices. The reader may find it necessary to check that all the concepts (incidence matrix, cycle and cut spaces, fundamental cycles and cuts) can be defined as before and the proofs of the results remain unchanged.

By considering multigraphs one can set up Theorem 7 in a slightly simpler form without adding a new vertex to the graph G' of the network. Thus if the current enters G' at a vertex and leaves it at a vertex b, then we join a to b by a *new edge* e of 0 conductance (even if a and b had been joined before), and postulate (by choosing $\mathbf{g} = (0, 0, \ldots, 0, 1)$, where e is the last edge) that the potential difference in e is 1. Using this set-up one can check that the ratio of the current in e_i to the total current (that is the current in e) is indeed given by Theorem 1 of §1, though this checking is rather tedious and involved. On the other hand, it is very easy to express the total number of spanning trees in a graph in terms of the incidence matrix.

To do this we make use of two facts. The first is an extension of the observation that B_T is invertible: the modulus of the determinant of an $(n-1) \times (n-1)$ submatrix of $\tilde{B}$ is 1 if the edges corresponding to the columns form a tree and 0 otherwise (cf. Exercise 19). The second is the Cauchy-Binet formula of linear algebra, stating that if K is a $p \times q$ matrix ($p \leq q$) and L is a $q \times p$ matrix then $\det KL = \sum_P \det K_P L_P = \sum_P \det K_P \det L_P$. Here the summation is over all p-subsets P of $\{1, 2, \ldots, q\}$, and K_P is the $p \times p$ submatrix of K formed by the columns of K indexed with elements of P and L_P is the submatrix of L formed by the corresponding rows of L. Putting the two together we arrive at the following formula.

Theorem 8. *The number of spanning trees of G is* $\det \tilde{B}\tilde{B}^t$. □

EXERCISES

1.⁻ Verify the formulae in the star-delta transformation.

In Exercises 2–6 all edges are assumed to have unit resistance.

2.⁻ Calculate the resistance of the network shown in Figure I.1 measured between the vertices 2 and 3.

3. For each different pair of vertices of a cube calculate the resistance between them.

4. What is the resistance between two adjacent vertices of (a) an octahedron, (b) a dodecahedron and (c) an icosahedron?

5. Suppose each edge of a connected network is in the same number of spanning trees. Prove that the total resistance between two adjacent vertices is $(n - 1)/e$, where n is the order and e is the size of the network. Verify your answers to Exercise 4.

6. By applying suitable star-delta transformations, calculate the resistance of a dodecahedron between the midpoints of two adjacent edges.

7. Give a detailed proof of Theorem 2.

8. Find the squared rectangle indicated in Figure II.11.

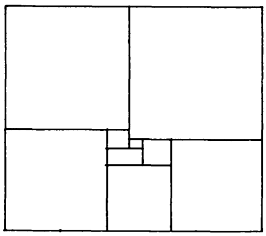


Figure II.11. A squaring of the 69×61 rectangle.

9. Which squared rectangle corresponds to the network in Figure II.12? Rotate the rectangle through 90° and draw the network for this rectangle.

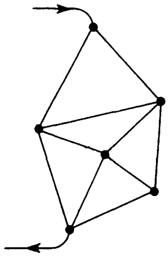


Figure II.12. A plane network.

10.⁻ How many essentially different squared rectangles correspond to the network of the cube in Exercise 3?

11. Find the tiling associated with Figure II.13.

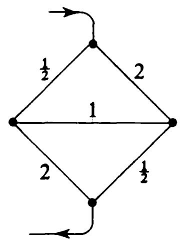


Figure II.13. A network with edges of differing resistance.

12. Find the simple perfect squared square given by the network in Figure II.14. (This example was found in 1964 by J. Wilson.)

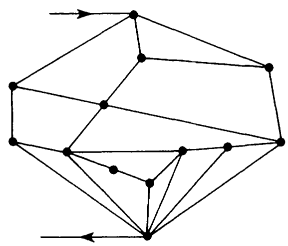


Figure II.14. A network giving a perfect square.

13. Show that an equilateral triangle cannot be dissected into finitely many incongruent equilateral triangles.

14. Prove that if a rectangular parallelepiped can be decomposed into cubes then the ratios of its sides are rational.

15. Show that a cube cannot be dissected into finitely many incongruent cubes.

16. Show that in a plane graph the boundaries of bounded faces form a cycle basis.

17. Show that the cycle space is the kernel of the map $C_1(G) \to C_0(G)$ defined by the incidence matrix B.

18. Let B^t be the transpose of B. Show that the cut space is the image of the map $C_0(G) \to C_1(G)$ defined by B^t.

19. Let F be a set of $n - 1$ edges of a graph of order n with incidence matrix B. Let $\bar{B}_F$ be an $(n - 1) \times (n - 1)$ submatrix of B whose columns correspond to the edges of F. Prove that $\bar{B}_F$ is invertible iff F is the edge set of a tree.

20. Deduce from Theorem 8 that there are n^{n-2} trees on n distinguishable vertices.

Notes

The first squared square was published by R. Sprague, Beispiel einer Zerlegung des Quadrats in lauter verschiedene Quadrate, *J. Reine Angew. Math.* **182** (1940) 60–64, closely followed by R. L. Brooks, C. A. B. Smith, A. H. Stone and W. T. Tutte, The dissection of rectangles into squares, *Duke Math. J.* **7** (1940) 312–340. The square shown in Fig. 8 was published in A. J. W. Duijvestijn, Simple perfect square of lowest order, *J. Combinational Theory Ser. B* **25** (1978) 240–243.

The original proof of Theorem 4 is due to Max Dehn, Über die Zerlegung von Rechtecken in Rechtecke, *Math. Ann.* **57** (1903) 314–322. Two survey articles the reader may wish to look at are W. T. Tutte, The quest of the perfect square, *Amer. Math. Monthly* **72** (1965), 29–35 and N. D. Kazarinoff and R. Weitzenkamp, Squaring rectangles and squares, *Amer. Math. Monthly* **80** (1973) 877–888.

CHAPTER III
Flows, Connectivity and Matching

The results of this chapter have many applications in diverse branches in mathematics; in particular, Hall's marriage theorem is a useful tool in algebra and analysis. Every result we are going to present is a necessary and sufficient condition for the existence of certain objects: in each case the beauty of the theorem is that a condition whose necessity is obvious is shown to be also sufficient. In the natural formulation of our results we shall have two functions, say f and g, clearly satisfying $f \leq g$, and we shall show that $\max f = \min g$. The results of the chapter are closely interrelated so the order they are proved in is a matter of taste; to emphasize this, some results will be given several proofs.

In the first chapter we looked at the simplest properties of connectivity. Now we shall go a step further and in §2 we shall prove the basic theorem about connectivity, namely Menger's theorem, first proved in 1927. This theorem turns out to be one of the many consequences of a fundamental theorem about flows in directed graphs, the max-flow min-cut theorem.

Though this book is devoted almost exclusively to undirected graphs, we introduce in the first paragraph the concept of a flow in a directed graph and go on to prove the max-flow min-cut theorem, partly because it provides one of the simplest proofs of Menger's theorem and partly because it has a number of other important consequences concerning undirected graphs.

A set of independent edges in a graph is called a *matching*. In a bipartite graph a matching can be identified with a flow and the max-flow min-cut theorem takes the form of another widely applicable combinatorial result, Hall's marriage theorem. We discuss various forms of this result in §3.

Though we deduce the theorems of both Menger and Hall from the max-flow min-cut theorem, the results are in fact very closely related, and

because of their fundamental importance we give also independent proofs of each.

Hall's theorem tells us, in particular, when a bipartite graph has got a 1-*factor*, a subgraph whose vertex set is that of the original graph and in which every vertex has degree 1. The question of the existence of a 1-factor in an arbitrary graph is considerably harder. It is answered by a theorem of Tutte that we present in §4.

§1 Flows in Directed Graphs

Let $\vec{G}$ be a (finite) directed graph with vertex set V and edge set $\vec{E}$. We shall study (static) flows in G from a vertex s (the *source*) to a vertex t (the *sink*). A *flow* f is a non-negative function defined on the edges; the value $f(\overrightarrow{xy})$ is the *flow* or *current* in the edge $\overrightarrow{xy}$. For notational simplicity we shall write $f(x, y)$ instead of $f(\overrightarrow{xy})$ and a similar convention will be used for other functions. The only condition a flow from s to t has to satisfy is Kirchhoff's current law: *the total current flowing into each intermediate vertex* (that is vertex different from s and t) *is equal to the total current leaving the vertex.* Thus if for $x \in V$ we put

$$\Gamma^+(x) = \{y \in V : \overrightarrow{xy} \in \vec{E}\},$$
$$\Gamma^-(x) = \{y \in V : \overrightarrow{yx} \in \vec{E}\},$$

then a *flow from s to t* satisfies the following condition:

$$\sum_{y \in \Gamma^+(x)} f(x, y) = \sum_{z \in \Gamma^-(x)} f(z, x) \quad \text{for each } x \in V - \{s, t\}.$$

Since

$$0 = \sum_{x \in V - \{s,t\}} \left\{ \sum_{y \in \Gamma^+(x)} f(x, y) - \sum_{z \in \Gamma^-(x)} f(z, x) \right\}$$
$$= \sum_{u \in \{s,t\}} \left\{ \sum_{z \in \Gamma^-(u)} f(z, u) - \sum_{y \in \Gamma^+(u)} f(u, z) \right\}$$

we find that

$$\sum_{y \in \Gamma^+(s)} f(s, y) - \sum_{y \in \Gamma^-(s)} f(y, s) = \sum_{y \in \Gamma^-(t)} f(y, t) - \sum_{y \in \Gamma^+(t)} f(t, y).$$

The common value, denoted by $v(f)$, is called the *value of* f or the *amount of flow* from s to t.

We wish to determine the maximal flow value from s to t provided the flow satisfies certain constraints. First we shall deal with the case when the

so called *capacity* of an edge restricts the current through the edge. It will turn out that several other seemingly more complicated restrictions can be reduced to this case.

Let us fix our directed graph $\vec{G} = (V, \vec{E})$ and two vertices in it, say s and t. With each edge $\overline{xy}$ of G we associate a non-negative number $c(x, y)$, called the *capacity* of the edge. We shall assume that the current flowing through the edge $\overline{xy}$ cannot be more than the capacity $c(x, y)$.

Given two subsets X, Y of V we write $\vec{E}(X, Y)$ for the set of directed X-Y edges:

$$\vec{E}(X, Y) = \{\overline{xy} : x \in X, y \in Y\}.$$

Whenever $g: \vec{E} \to \mathbb{R}$ is a function, we put

$$g(X, Y) = \sum g(x, y),$$

where the summation is over $\vec{E}(X, Y)$. If S is a subset of V containing s but not t then $\vec{E}(S, \bar{S})$ is called a *cut* separating s from t. (Here $\bar{S} = V - S$.) If we delete the edges of a cut then no positive-valued flow from s to t can be defined on the remainder. Conversely, it is easily seen that if F is a set of edges after whose deletion there is no flow from s to t (that is $v(f) = 0$ for every flow from s to t) then F contains a cut (Exercise 1). The *capacity* of a cut $\vec{E}(S, \bar{S})$ is $c(S, \bar{S})$ (see Figure III.1). It is easily seen (Exercise 2) that the capacity of a cut is at least as large as the value of any flow, so the minimum of all cut capacities is at least as large as the maximum of all flow values. The celebrated max-flow min-cut theorem of Ford and Fulkerson states that this trivial inequality is, in fact, an equality. Before stating this theorem and getting down to the proof, let us justify the above use of the words minimum and maximum. Since there are only finitely many cuts, there is a cut whose capacity is minimal. The existence of a flow with maximal value is only slightly less trivial. Indeed,

$$v(f) \le \sum_{\overline{xy} \in \vec{E}} c(x, y)$$

for every flow f, so $v = \sup v(f) < \infty$. Let $f_1, f_2, \ldots$ be a sequence of flows with $\lim_n v(f_n) = v$. Then, by passing to a subsequence, we may assume that for each $\overline{xy} \in \vec{E}$ the sequence $(f_n(x, y))$ is convergent, say to $f(x, y)$. The

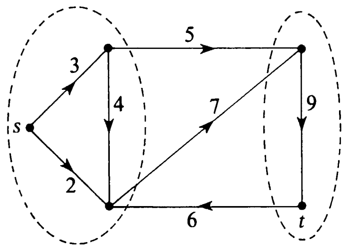


Figure III.1. A cut with capacity 12. (The numbers next to the edges indicate their capacity.)

function f is a flow with value v, that is a flow with maximal value. In a similar way one can show that even if some of the edges have infinite capacity, there is a flow with maximal value which can be either finite or infinite (Exercise 3).

Theorem 1. (Max-flow min-cut theorem.) *The maximal flow value from s to t is equal to the minimum of the capacities of cuts separating s from t.*

PROOF. We have remarked already that there is a flow f with maximal value, say v, and the capacity of every cut is at least v. Thus, in order to prove the theorem we have to show that there is a cut with capacity v. We shall, in fact, do considerably more than this: we shall give a very simple procedure for constructing such a cut from a flow f with maximal value.

Define a subset $S \subset V$ recursively as follows. Let $s \in S$. If $x \in S$ and

$$c(x, y) > f(x, y)$$

or

$$f(y, x) > 0$$

then let $y \in S$.

We claim that $\vec{E}(S, \bar{S})$ is a cut separating s from t with capacity $v = v(f)$. Let us see first why t cannot belong to S. If t belongs to S, we can find vertices $x_0 = s, x_1, \ldots, x_l = t$ such that

$$\varepsilon_i = \max\{c(x_i, x_{i+1}) - f(x_i, x_{i+1}), f(x_{i+1}, x_i)\} > 0$$

for every i, $0 \leq i \leq l - 1$. Put $\varepsilon = \min_i \varepsilon_i$. Then f can be augmented to a flow f^* in the following way: if $\varepsilon_i > f(x_{i+1}, x_i)$ then increase the flow in $\overrightarrow{x_i x_{i+1}}$ by ε, otherwise decrease the flow in $\overrightarrow{x_{i+1} x_i}$ by ε. Clearly f^* is a flow and its value is $v(f^*) = v(f) + \varepsilon$, contradicting the maximality of f. This shows that $t \notin S$ so $\vec{E}(S, \bar{S})$ is a cut separating s from t.

Now $v(f)$ is equal to the value of the flow from S to $\bar{S}$ defined in the obvious way:

$$\sum_{x \in S, y \in \bar{S}} f(x, y) - \sum_{x \in \bar{S}, y \in S} f(x, y).$$

By the definition of S the first sum is exactly

$$\sum_{x \in S, y \in \bar{S}} c(x, y) = c(S, \bar{S})$$

and each summand in the second sum is zero. Hence $c(S, \bar{S}) = v(f)$, as required. □

The max-flow min-cut theorem is the cornerstone of the theory to be presented in this chapter. Note that the theorem remains valid (with exactly the same proof) if some of the edges have infinite capacity but the maximal flow value is finite.

The proof of the theorem also provides a surprisingly efficient *algorithm* for finding a flow with maximal value if the capacity function is *integral*,

that is if $c(x, y)$ is an integer for every edge $\overrightarrow{xy}$. We start with the identically zero flow: $f_0(x, y) = 0$ for every $\overrightarrow{xy} \in \vec{E}$. We shall construct an increasing sequence of flows $f_0, f_1, f_2, \ldots$ that has to terminate in a maximal flow. Suppose we have constructed f_i. As in the proof above, we find the set S belonging to f_i. Now if $t \notin S$ then f_i is a maximal flow (and $E(S, \bar{S})$ is a minimal cut) so we terminate the sequence. If, on the other hand, $t \in S$, then f_i can be augmented to a flow f_{i+1} by increasing the flow along a path from s to t. Since each $v(f_i)$ is an integer, we have $v(f_i) \geq v(f_{i-1}) + 1$, and the sequence must end in at most $\sum_{x,y} c(x, y)$ steps.

Moreover, if c is integral the algorithm constructs a maximal flow which is also integral, that is a flow whose value is an integer in every edge. Indeed, f_0 is integral and if f_{i-1} is integral then so is f_i, since it is obtained from f_{i-1} by increasing the flow in a path by a value which is the minimum of positive integers. This result is often called the integrality theorem.

Theorem 2. *If the capacity function is integral then there is a maximal flow which is also integral.* □

We shall rely on this simple result when we use the max-flow min-cut theorem to find various paths in graphs. It is important to note that the results do not claim uniqueness: the algorithm finds one of the maximal flows (usually there are many) and Theorem 2 claims that one of the maximal flows is integral.

The existence of the algorithm proves some other intuitively obvious results as well. For instance, there is a maximal *acyclic* flow, that is one which does not contain a flow around a cycle $x_1 x_2, \ldots, x_k$:

$$f(x_1, x_2) > 0, f(x_2, x_3) > 0, \ldots, f(x_{k-1}, x_k) > 0, f(x_k, x_1) > 0.$$

It may seem somewhat surprising that if instead of one source and one sink we take several of each, the situation does not become any more complicated. The only occasion we have to be a little more careful is the definition of a cut. If $s_1, \ldots, s_k$ are the sources and $t_1, \ldots, t_l$ are the sinks then $\vec{E}(S, \bar{S})$ is a *cut* if $s_i \in S$ and $t_j \in \bar{S}$ for every i, j, $1 \leq i \leq k$, $1 \leq j \leq l$.

In order to be able to apply the max-flow min-cut theorem, let us add a new source s and a new sink t to $\vec{G}$, together with all the edges $\overrightarrow{ss_i}$ and $\overrightarrow{t_j t}$ of *infinite capacity*. Let $\vec{H}$ be the graph obtained in this way. Consider those flows from $s_1, \ldots, s_k$ to $t_1, \ldots, t_l$ in $\vec{G}$ in which the total current entering (leaving) a source (sink) is not greater than the total current leaving (entering) it. These flows can easily be extended to a flow from s to t in $\vec{H}$ and this extension establishes a 1-1 correspondence between the two sets of flows. Furthermore, a cut separating s from t in $\vec{H}$ that has finite capacity cannot contain an edge of the form $\overrightarrow{ss_i}$ and $\overrightarrow{t_j t}$, so it corresponds to a cut of the same capacity in $\vec{G}$, separating $s_1, \ldots, s_k$ from $t_1, \ldots, t_l$. Thus Theorem 1 has the following extension.

Theorem 3. *The maximum of the flow value from a set of sources to a set of sinks is equal to the minimum of the capacity of cuts separating the sources from the sinks.* □

Let us assume now that we have capacity restrictions on the *vertices*, except the source and the sink. Thus we are given a function $c: V - \{s, t\} \to \mathbb{R}^+$ and every flow f from s to t has to satisfy the following inequality:

$$\sum_{y \in \Gamma^+(x)} f(x, y) = \sum_{z \in \Gamma^-(x)} f(z, x) \leq c(x) \quad \text{for every } x \in V - \{s, t\}.$$

How should we define a cut in this situation? A *cut* is a subset S of $V - \{s, t\}$ such that no positive-valued flow from s to t can be defined on G-S. (In order to distinguish the two kinds of cuts we sometimes call this a vertex-cut and the other one an edge-cut. However, it is almost always clear which cut is in question.) Can we carry over the max-flow min-cut theorem to this case? Yes, very easily, if we notice that a flow can be interpreted to flow in a vetex as well, namely from the part where all the currents enter it to the part where all the currents leave it. More precisely we can turn each vertex of $\vec{G}$ into an edge (without changing the nature of the directed graph) in such a way that any current entering (and leaving) the vertex will be forced through the edge. To do this, replace each vertex $x \in V - \{s, t\}$ by two vertices, say x_- and x_+, send each incoming edge to x_- and send each outgoing edge to x_+. Finally, add edges from x_- to x_+ with capacity $c(x_-, x_+) = c(x)$ (see Figure III.2). There is a simple 1-1 correspondence between the flows from

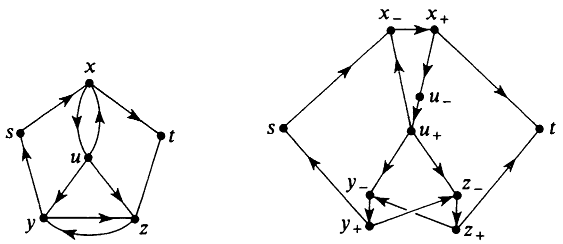


Figure III.2. Replacing a graph $\vec{G}$ with restrictions on the capacity of the vertices by a graph $\vec{H}$ with restrictions on the capacity of the edges.

s to t in $\vec{G}$ satisfying the capacity restrictions on the vertices and the flows in the new graph $\vec{H}$ satisfying the capacity restrictions on (some of) the edges. Since in $\vec{H}$ only the edges $\overrightarrow{x_- x_+}$ have finite capacities, an edge-cut of finite capacity in $\vec{H}$ consists entirely of edges of the form $\overrightarrow{x_- x_+}$, so it corresponds to a vertex-cut in $\vec{G}$ of the same capacity. Thus we have the following form of Theorem 1.

Theorem 4. *Let $\vec{G}$ be a directed graph with capacity bounds on the vertices other than the source s and the sink t. Then the minimum of the capacity of a vertex-cut is equal to the maximum of the flow value from s to t.*

Theorems 1, 3 and 4 can easily be combined into a single theorem. We leave this to the reader (Exercise 6).

§2 Connectivity and Menger's Theorem

Recall that if any two vertices of a graph can be joined by a path then the graph is said to be *connected*, otherwise it is *disconnected*. A maximal connected subgraph of a graph G is a *component* of G.

If G is connected and G-W is disconnected, where W is a set of vertices or a set of edges, then we say that W *separates* G. If in G-W two vertices s and t belong to different components then W *separates* s *from* t. We say that a graph G is *k-connected* ($k \geq 2$) if either G is aK^{k+1} or else it has at least $k + 2$ vertices and no set of $k - 1$ vertices separates it. Similarly G is *k-edge-connected* ($k \geq 2$) if it has at least two vertices and no set of at most $k - 1$ edges separates it. A connected graph is also said to be 1-*connected* and 1-*edge-connected*. The maximal value of k for which a connected graph G is k-connected is the *connectivity* of G, denoted by $\kappa(G)$. (If G is disconnected, we put $\kappa(G) = 0$.) The *edge-connectivity* $\lambda(G)$ is defined analogously.

Clearly a graph is 2-connected iff it is connected, has at least 3 vertices and contains no cutvertex. Similarly, a graph is 2-edge-connected iff it is connected, has at least 2 vertices and contains no bridge. It is often easy to determine the connectivity of a given graph. Thus if $1 \leq l \leq n$ then $\kappa(P^l) = \lambda(P^l) = 1$, $\kappa(C^n) = \lambda(C^n) = 2$, $\kappa(K^n) = \lambda(K^n) = n - 1$ and $\kappa(K^{l,n}) = \lambda(K^{l,n}) = l$. In order to correct the false impression that the vertex-connectivity is equal to the edge-connectivity, note that if G is obtained from the disjoint union of two K^l by adding a new vertex x and joining x to every old vertex, then $\kappa(G) = 1$, since x is a cutvertex, but $\lambda(G) = l$ (see also Exercise 11). This last example shows that $\lambda(G - x)$ may be 0 even when $\lambda(G)$ is large. However, it is clear from the definitions that for every vertex x and edge xy we have

$$\kappa(G) - 1 \leq \kappa(G - x) \quad \text{and} \quad \lambda(G) - 1 \leq \lambda(G - xy) \leq \lambda(G).$$

If G is non-trivial (that is has at least two vertices), then the parameters $\delta(G)$, $\lambda(G)$ and $\kappa(G)$ satisfy the following inequality:

$$\kappa(G) \leq \lambda(G) \leq \delta(G).$$

Indeed, if we delete all the edges incident with a vertex, the graph becomes disconnected, so the second inequality holds. To see the other inequality, note first that if G is complete then $\kappa(G) = \lambda(G) = |G| - 1$, and if $\kappa(G) \leq 1$ then $\lambda(G) = \kappa(G)$. Suppose now that G is not complete, $\lambda(G) = k \geq 2$ and $\{x_1 y_1, x_2 y_2, \ldots, x_k y_k\}$ is a set of edges disconnecting G. If $G - \{x_1, x_2, \ldots, x_k\}$ is disconnected then $\kappa(G) \leq k$. Otherwise each vertex x_i

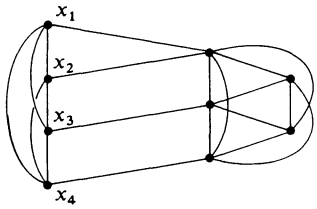


Figure III.3. A 4-edge-connected graph G such that $G - x_1, x_2, x_3, x_4$ is connected.

has degree at most k (and so exactly k), as shown in Figure III.3. Deleting the neighbours of x_1, we disconnect G. Hence $\kappa = \lambda(G)$.

Another property immediate from the definition of vertex-connectivity is that if G_1 and G_2 are k-connected ($k \geq 1$) subgraphs of a graph G having at least k common vertices, then $G_1 \cup G_2$ is also k-connected. Indeed if $W \subset V(G_1) \cup V(G_2)$ has at most $k - 1$ vertices, then there is a vertex x in $V(G_1) \cap V(G_2)\backslash W$. Therefore the connected subgraphs $G_1 - W$ and $G_2 - W$ of G have at least one vertex, namely x, in common, so $G_1 \cup G_2 - W = (G_1 - W) \cup (G_2 - W)$ is connected.

Having seen in Chapter I how useful it is to partition a graph into its components, that is into its maximal connected subgraphs, let us attempt a similar decomposition using all maximal 2-connected subgraphs. A subgraph B of a graph G is a *block of G* if either B is a bridge (and its endvertices) or else it is a maximal 2-connected subgraph of G. The remarks above show that any two blocks have at most one vertex in common and if x, y are distinct vertices of a block B then $G - E(B)$ contains no x-y path. Therefore every vertex belonging to two blocks is a cutvertex of G, and, conversely, every cutvertex belongs to at least two blocks. Recalling that a cycle is 2-connected and an edge is a bridge iff no cycle contains it, we find that G decomposes into its blocks $B_1, B_2, \ldots, B_p$ in the following sense:

$$E(G) = \bigcup_1^p E(B_i) \quad \text{and} \quad E(B_i) \cap E(B_j) = \emptyset \quad \text{if } i \neq j.$$

Suppose now that G is a non-trivial connected graph. Let $bc(G)$ be the graph whose vertices are the blocks and cutvertices of G and whose edges join cutvertices to blocks: each cutvertex is joined to the blocks containing it. Then $bc(G)$, called the block-cutvertex graph of G, is a tree. Each endvertex of $bc(G)$ is a block of G, called an *endblock* of G. If G is 2-connected or it is a K^2 (an "edge") then it contains only one block, itself; otherwise there are at least two endblocks, and a block is an endblock iff it contains exactly one cutvertex (Figure III.4).

The basic result in the theory of connectivity was proved by Menger in 1927. It is the analogue of the max-flow min-cut theorem for (non-directed) graphs. Recall that two s-t paths are *independent* if they have only the vertices s and t in common.

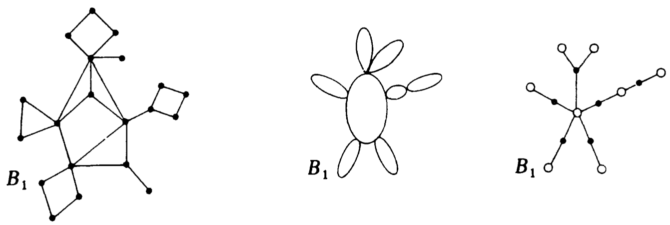


Figure III.4. The construction of the block-cutvertex tree $bc(G)$. B_1 is an endblock.

Theorem 5.

i. *Let s and t be distinct non-adjacent vertices of a graph G. Then the minimal number of vertices separating s from t is equal to the maximal number of independent s-t paths.*
ii. *Let s and t be distinct vertices of G. Then the minimal number of edges separating s from t is equal to the maximal number of edge-disjoint s-t paths.*

PROOF. (i) Replace each edge xy of G by two *directed* edges, $\vec{xy}$ and $\vec{yx}$, and give each vertex other than s and t capacity 1. Then by Theorem 4 the maximal flow value from s to t is equal to the minimum of the capacity of a cut separating s from t. By the integrality theorem (Theorem 2) there is a maximal flow with current 1 or 0 in each edge. Therefore the maximum flow value from s to t is equal to the maximal number of independent s-t paths. The minimum of the cut capacity is clearly the minimal number of vertices separating s from t.

(ii) Proceed as in (i), except instead of restricting the capacity of the vertices, give each directed edge capacity 1. □

The two parts of the above theorem are called the vertex form of Menger's theorem and the edge form of Menger's theorem. One can easily deduce the edge form from the vertex form (Exercise 14), but the other implication is not so easily proved. Since, as we have mentioned already, the max-flow min-cut theorem can also be deduced from Menger's theorem, we shall give another proof of the vertex form of Menger's theorem from first principles.

SECOND PROOF OF THE VERTEX FORM OF MENGER'S THEOREM. Denote by k the minimal number of vertices separating s and t. Then clearly there are at most k independent s-t paths and for $k \leqq 1$ *there are* k independent s-t paths.

Suppose the theorem fails. Take the minimal $k \geq 2$ for which there is a counterexample to the theorem and let G be a counterexample (for the minimal k) with the minimal number of edges. Then there are at most $k - 1$ independent s-t paths and no vertex x is joined to both s and t, otherwise $G - x$ would be a counterexample for $k - 1$.

Let W be a set of k vertices separating s from t. Suppose neither s nor t is adjacent to every vertex in W. Let G_s be obtained from G by replacing the component of $G - W$ containing s by a single vertex s' and joining s' to each vertex in W. In G_s we still need k vertices to separate s' from t and since the component we collapsed had at least two vertices, G_s has fewer edges than G. Since G is a counterexample of minimal size, in G_s there are k independent s'-t paths. The segments of these k paths from t to W are such that any two of them have exactly the vertex t in common. In particular, for every $w \in W$ one of these paths is a t-w path. If we carry out the analogous procedure for t instead of s then we get k paths from s to W. These two sets of paths can be put together to give k independent s-t paths, contradicting our assumption. Hence for any set W of k vertices separating s from t *either s or t is adjacent to each vertex of W.*

Let $sx_1x_2 \dots x_lt$ be a shortest s-t path. Then $l \geq 2$ and, by the minimality of G, in the graph $G - x_1x_2$ we can find a set W_0 of $k - 1$ vertices separating s from t. Then both $W_1 = \{x_1\} \cup W_0$ and $W_2 = \{x_2\} \cup W_0$ are k-sets separating s from t. Since t is not joined to x_1, s is joined to every vertex in W_1. Similarly, s is not joined to x_2, so t is joined to every vertex in W_2. This implies the contradiction that s and t have at least one common neighbour: every vertex in W_0 is a common neighbour of s and t and $|W_0| = k - 1 \geq 1$. □

Corollary 6. *A graph is k-connected ($k \geq 2$) iff it has at least two vertices and any two vertices can be joined by k independent paths. A graph is k-edge-connected ($k \geq 2$) iff it has at least two vertices and any two vertices can be joined by k edge-disjoint paths.*

Another characterization of k-connectivity is given in Exercise 12.

Corresponding to the max-flow min-cut theorem for multiple sources and sinks, one has the following version of Menger's theorem. If S and T are arbitrary subsets of vertices of G then the maximal number of vertex-disjoint (including endvertices!) S-T paths is $\min\{|W| : W \subset V(G), G - W$ has no S-T path$\}$. To see this add two new vertices to G, say s and t, join s to every vertex in S and t to every vertex in T, and apply Menger's theorem to the vertices s and t in the new graph.

§3 Matching

Given a finite group G and a subgroup H of index m, can you find m elements of G, say $g_1, g_2, \dots, g_m$, such that $\{g_1H, g_2H, \dots, g_mH\}$ is the set of left cosets of H and $\{Hg_1, Hg_2, \dots, Hg_m\}$ is the set of right cosets? A reformulation of this problem turns out to be a special case of the following problem arising frequently in diverse branches of mathematics. Given a family

$\mathscr{A} = \{A_1, A_2, \ldots, A_m\}$ of subsets of a set X, can we find m distinct elements of X, one from each A_i? A set $\{x_1, x_2, \ldots, x_m\}$ with these properties (i.e. $x_i \in A_i$, $x_i \neq x_j$ if $i \neq j$) is called a *set of distinct representatives* of the family $\mathscr{A}$. The set system $\mathscr{A}$ is naturally identifiable with a bipartite graph with vertex classes $V_1 = \mathscr{A}$ and $V_2 = X$ in which $A_i \in \mathscr{A}$ is joined to every $x \in X$ contained in A_i. A system of distinct representatives is then a *set of m independent edges* (thus each vertex in V_1 is incident with one of these edges). We also say that there is a *complete matching from V_1 to V_2*.

It is customary to formulate this problem in terms of marriage arrangements. Given m girls and n boys, under what conditions can we marry off all the girls provided we do not want to carry matchmaking as far as to marry a girl to a boy she does not even know?

It is clear that both the max-flow min-cut theorem and Menger's theorem imply a necessary and sufficient condition for the existence of a complete matching. In fact, because of the special features of a bipartite graph, there is a particularly simple and pleasant necessary and sufficient condition.

If there are k girls who altogether know at most $k - 1$ boys then we cannot find suitable marriages for these girls. Equivalently, if there is a complete matching from V_1 to V_2 then for every $S \subset V_1$ there are at least $|S|$ vertices of V_2 adjacent to a vertex in S, that is

$$|\Gamma(S)| \geq |S|.$$

The result that this necessary condition is also sufficient is usually called *Hall's theorem.* (P. Hall proved it in 1935, and an equivalent form of it was proved by König and Egerváry in 1931, but both versions follow immediately from Menger's theorem, proved in 1927.) We shall give three proofs. The first is based on Menger's theorem or the max-flow min-cut theorem, the other two prove the result from first principles.

Theorem 7. *A bipartite graph G with vertex sets V_1 and V_2 contains a complete matching from V_1 to V_2 iff*

$$|\Gamma(S)| \geq |S| \quad \textit{for every } S \subset V_1.$$

We have already seen that the condition is necessary so we have to prove only the sufficiency.

1ST PROOF. Both Menger's theorem (applied to the *sets* V_1 and V_2, as at the end of §2) and the max-flow min-cut theorem (applied to the directed graph obtained from G by sending each edge from V_1 to V_2, in which each vertex has capacity 1) imply the following. If G does not contain a complete matching from V_1 to V_2 then there are $T_1 \subset V_1$ and $T_2 \subset V_2$ such that $|T_1| + |T_2| < |V_1|$ and there is no edge from $V_1 - T_1$ to $V_2 - T_2$. Then $\Gamma(V_1 - T_1) \subset T_2$ so

$$|\Gamma(V_1 - T_1)| \leq |T_2| < |V_1| - |T_1| = |V_1 - T_1|.$$

This shows the sufficiency of the condition. □

2ND PROOF. In this proof, due to Halmos and Vaughan, we shall use the matchmaking terminology. We shall apply induction on m, the number of girls. For $m = 1$ the condition is clearly sufficient so we assume that $m \geq 2$ and the condition is sufficient for smaller values of m.

Suppose first that any k girls ($1 \leq k < m$) know at least $k + 1$ boys. Then we arrange one marriage arbitrarily. The remaining sets of girls and boys still satisfy the condition, so the other $m - 1$ girls can be married off by induction.

Suppose now that for some k, there are k girls who altogether know exactly k boys. These girls can clearly be married off by induction. What about the other girls? We can marry them off (again by induction) if they also satisfy the condition provided we do not count the boys who are already married. But the condition is satisfied since if some l girls to be married know fewer than l remaining boys then these girls together with the first k girls would know fewer than $k + l$ boys. □

3RD PROOF. This proof is due to Rado. Let G be a minimal graph satisfying the condition. It suffices to show that G consists of $|V_1|$ *independent edges*.

If this is not so then G contains two edges of the form $a_1 x, a_2 x$, where $a_1, a_2 \in V_1$ and $x \in V_2$. Since the deletion of either of these edges invalidates the condition, there are sets $A_1, A_2 \subset V_1$ such that $|\Gamma(A_i)| = |A_i|$ and a_i is the only vertex of $A_i (i = 1, 2)$ adjacent to x. Then

$$\begin{aligned} |\Gamma(A_1) \cap \Gamma(A_2)| &= |\Gamma(A_1 - \{a_1\}) \cap \Gamma(A_2 - \{a_2\})| + 1 \\ &\geq |\Gamma(A_1 \cap A_2)| + 1 \geq |A_1 \cap A_2| + 1, \end{aligned}$$

which implies the following contradiction:

$$\begin{aligned} |\Gamma(A_1 \cup A_2)| = |\Gamma(A_1) \cup \Gamma(A_2)| &= |\Gamma(A_1)| + |\Gamma(A_2)| \\ &- |\Gamma(A_1) \cap \Gamma(A_2)| \leq |A_1| + |A_2| - |A_1 \cap A_2| - 1. \end{aligned}$$

A regular bipartite graph satisfies the conditions of Hall's theorem, so it has a complete matching. In turn this implies that we can indeed find group elements $g_1, g_2, \ldots, g_m$ as required at the beginning of the section.

Let us reformulate the marriage theorem in terms of sets of distinct representatives.

Theorem 7'. *A family $\mathscr{A} = \{A_1, A_2, \ldots, A_m\}$ of sets has a set of distinct representatives iff*

$$\left| \bigcup_{i \in F} A_i \right| \geq |F| \quad \textit{for every } F \subset \{1, 2, \ldots, m\}. \qquad \square$$

There are two natural extensions of the marriage theorem. Suppose the marriage condition is not satisfied. How near can we come to marrying off

all the girls? When can we marry off all the girls but d of them? Clearly only if any k of them know at least $k - d$ boys. This obvious necessary condition is again sufficient.

Corollary 8. *Suppose the bipartite graph $G = G_2(m, n)$ with vertex sets V_1, V_2 satisfies the following condition:*

$$|\Gamma(S)| \geq |S| - d \quad \textit{for every } S \subset V_1.$$

Then G contains $m - d$ independent edges.

PROOF. Add d vertices to V_2 and join them to each vertex in V_1. The new graph G^* satisfies the condition for a complete matching. At least $m - d$ of the edges in this matching belong to G. □

The next extension concerns matchmaking *for boys* in a polygynous country, where the ith boy intends to marry d_i girls.

Corollary 9. *Let G be a bipartite graph with vertex classes $V_1 = \{x_1, \ldots, x_m\}$ and $V_2 = \{y_1, \ldots, y_n\}$. Then G contains a subgraph H such that $d_H(x_i) = d_i$ and $0 \leq d_H(y_i) \leq 1$ iff*

$$|\Gamma(S)| \geq \sum_{x_i \in S} d_i \quad \textit{for every } S \subset V_1.$$

PROOF. Replace each vertex x_i by d_i vertices joined to each vertex in $\Gamma(x_i)$. Then there is a subgraph H iff the new graph has a matching from the new first vertex class to V_2. The result follows from Theorem 7. □

Of course, Corollary 9 also has a defect form which the reader is encouraged to state and deduce from this.

The reader is probably aware of the fact that these corollaries are still special cases of the max-flow min-cut theorem. In fact, the bipartite graph version of the max-flow min-cut theorem is considerably more general than the corollaries above.

Theorem 10. *Let $G = G_2(m, n)$ be a bipartite graph with vertex classes $V_1 = \{x_1, \ldots, x_m\}$ and $V_2 = \{y_1, \ldots, y_n\}$. For $S \subset V_1$ and $1 \leq j \leq n$ denote by S_j the number of edges from y_j to S. Let $d_1, \ldots, d_m$ and $e_1, \ldots, e_n$ be natural numbers and let $d \geq 0$. Then there exists a subgraph H of G with*

$$e(H) \geq \sum_1^n d_i - d,$$

$$d_H(x_i) \leq d_i, \qquad 1 \leq i \leq m$$

and

$$d_H(y_i) \leq e_j, \qquad 1 \leq j \leq n,$$

iff for every $S \subset V_1$ *we have*

$$\sum_{x_i \in S} d_i \leq \sum_{j=1}^{n} \min\{S_j, e_j\} + d.$$

PROOF. Turn G into a directed graph $\vec{G}$ by sending each edge from V_1 to V_2. Give each edge capacity 1, a vertex x_i capacity d_i and a vertex y_j capacity e_j. Then there is a subgraph H with required properties iff in $\vec{G}$ there is a flow from V_1 to V_2 with value at least $\sum_1^m d_i - d$ and, by the max-flow min-cut theorem, this happens iff every cut has capacity at least $\sum_1^m d_i - d$. Now minimal cut-sets are of the form $T \cup U \cup E(V_1 - T, V_2 - U)$, where $T \subset V_1$ and $U \subset V_2$. Given a set T, the capacity of such a cut will be minimal if a vertex y_j belongs to U iff its capacity is smaller than the number of edges from $S = V_1 - T$ to y_j. With this choice of U the capacity of the cut is exactly

$$\sum_{x_i \in T} d_i + \sum_{1}^{n} \min\{S_j, e_j\}.$$

The condition that this is at least $\sum_1^m d_i - d$ is clearly the condition in the theorem. □

The reader is invited to check that the second proof of Theorem 7 can be rewritten word for word to give a proof of the exact form of this result (that is with $d = 0$) and the defect form (the case $d \geq 0$) can be deduced from it as Corollary 8 was deduced from Theorem 7 (Exercise 26).

To conclude this section we prove another extension of the marriage theorem. This is *Dilworth's theorem* concerning partially ordered sets. A *partial order* $<$ on a set is a transitive and irreflexive relation defined on some ordered pairs of elements. Thus if $x < y$ and $y < x$ then $x < z$, but $x < y$ and $y < x$ cannot both hold. A set with a partial order on it is a *partially ordered set*. The relation $x \leq y$ expresses the fact that either $x = y$ or else $x < y$. A subset C of a partially ordered set P is a *chain* (or *tower*) if for $x, y \in C$ either $x \leq y$ or $y < x$. A set $A \subset P$ is an *antichain* if $x < y$ implies that $\{x, y\} \not\subset A$. See Figure III.5 for an example.

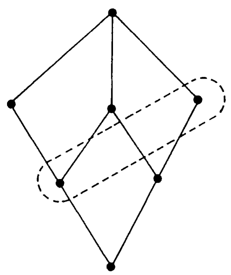


Figure III.5. A partially ordered set and a maximal antichain. (An edge indicates that its upper endvertex is greater than its lower endvertex.)

What is the smallest number of chains into which we can decompose a partially ordered set? Since no two elements of an antichain can belong to the same chain, we need at least as many chains as the minimal size of an antichain. Once again, the trivial necessary condition is, in fact, sufficient.

Theorem 11. *If every antichain in a (finite) partially ordered set P has at most m elements then P is the union of m chains.*

PROOF. Let us apply induction on $|P|$. If $P = \emptyset$ there is nothing to prove so we suppose that $|P| > 0$ and the theorem holds for sets with fewer elements.

Let C be a maximal chain in P. (Thus if $x \notin C$ then $C \cup \{x\}$ is no longer a chain.) If no antichain of $P - C$ has m elements then we are home by induction. Therefore we may assume that $P - C$ contains an antichain $A = \{a_1, a_2, \ldots, a_m\}$.

Define the *lower shadow* of A as

$$S^- = \{x \in P : x \leq a_i \text{ for some } i\},$$

and define the *upper shadow* S^+ of A analogously. Then P is the union of the two shadows since otherwise A could be extended to an antichain with $m + 1$ elements. Furthermore, neither shadow is the whole of P since the maximal element of C does not belong to S^- and the minimal element of C does not belong to S^+. By the induction hypothesis both shadows can be decomposed into m chains, say

$$S^- = \bigcup_1^m c_i^- \quad \text{and} \quad S^+ = \bigcup_1^m c_i^+.$$

Since different a_i belong to different chains, we may assume that $a_i \in C_i^-$ and $a_i \in C_i^+$.

The proof will be completed if we show that a_i is the maximal element of C_i^- and the minimal element of C_i^+. For in that case the chains C_i^- and C_i^+ can be strung together to give a single chain C_i and then $P = \bigcup_1^m C_i$.

Suppose then that, say, a_i is not the maximal element of C_i^-: $a_i < x$ for some $x \in C_i^-$. Since x is in the lower shadow of A, there is an $a_j \in A$ with $x \leq a_j$. However, this implies the contradiction $a_i < a_j$. □

§4 Tutte's 1-Factor Theorem

A *factor* of a graph is a subgraph whose vertex set is that of the whole graph. If every vertex of a factor has degree r then we call it an *r-factor*. How can we characterize graphs with a 1-factor? If G has a 1-factor H and we delete a set S of vertices of G, then in a component C of $G - S$ and even number of vertices are on edges of H contained in C and the other vertices of C are

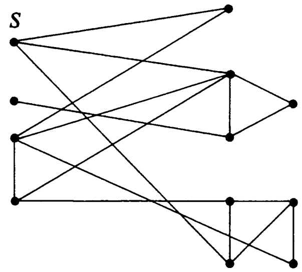


Figure III.6. A graph G with a 1-factor: $|S| = 4$ and $G - S$ has 2 odd components.

on edges of H joining a vertex of C to a vertex of S. In particular, for every *odd* component C of $G - S$ (that is a component with an odd number of vertices) there is an edge of H joining a vertex of C to a vertex of S. Now the edges of H are independent, so this implies that the graph $G - S$ has *at most* $|S|$ *odd components*, one for each vertex in S (see Figure III.6).

The necessity of the condition we have just found is rather trivial, but it is not clear at all that the condition is also sufficient. This surprising and deep result was first proved by Tutte in 1947. It will be convenient to denote by $q(H)$ the number of odd components of a graph H, that is the number of components of odd order.

Theorem 12. *A graph G has got a 1-factor iff*

$$q(G - S) \leq |S| \quad \textit{for every } S \subset V(G). \tag{*}$$

Proof. We know that the condition is necessary. We shall prove the sufficiency by induction on the order of G. For $|G| = 0$ there is nothing to prove. Now let G be a graph of order at least one satisfying (*) and suppose that the theorem holds for graphs of smaller order.

Let $S_0 \subset V(G)$ be a non-empty set for which equality holds in (*). Denote by $C_1, C_2, \ldots, C_m$, $m = |S_0| \geq 1$, the odd components of $G - S_0$ and let $D_1, D_2, \ldots, D_k$ be the even components of $G - S_0$. If the theorem is true and G does contain a 1-factor F then for each C_i there is at least one edge of F that joins a vertex of C_i to a vertex in S_0. Since $m = |S_0|$, for each C_i there is exactly one such edge, say $c_i s_i$, $c_i \in C_i$, $s_i \in S_0$. Each $C_i - c_i$ contains a 1-factor, (a subgraph of F), and each D_j contains a 1-factor (a subgraph of F). Finally, the edges $s_1c_1, s_2c_2, \ldots, s_mc_m$ form a complete matching from S_0 into the set $\{C_1, C_2, \ldots, C_m\}$.

The proof is based on the fact that one can find an S_0 which has all the properties described above. How shall we find such a set S_0? Let S_0 be a maximal non-empty subset of $V(G)$ for which equality holds in (*). Of course, a priori it is not even clear that there is such a set S_0. With $S = \varnothing$ the condition (*) implies that G has an even order. If s is any vertex of G then $G - \{s\}$ has even order so it has at least one odd component. Since (*)

holds, $G - \{s\}$ has exactly one odd component. Hence for every $S = \{s\}$ we have equality in (*). This establishes the existence of S_0.

As before, $C_1, C_2, \ldots, C_m$, $m = |S_0|$, are the odd components of $G - S_0$ and $D_1, D_2, \ldots, D_k$ are the even components.

(i) *Each D_j has a 1-factor.* Indeed, if $S \subset V(D_j)$ then

$$q(G - S_0 \cup S) = q(G - S_0) + q(D_j - S_0) \leq |S_0 \cup S| = |S_0| + |S|,$$

so

$$q(D_j - S) \leq |S|.$$

Hence by the induction hypothesis D_j has a 1-factor.

(ii) *If $c \in C_j$ then $C_i - c$ has a 1-factor.* Assume that this is false. Then by the induction hypothesis there is a set $S \subset V(C_i) - \{c\}$ such that

$$q(C_i - \{c\} \cup S) > |S|.$$

Since

$$q(C_i - \{c\} \cup S) + |S \cup \{c\}| \equiv |C_i| \equiv 1 \pmod 2,$$

this implies that

$$q(C_i - \{c\} \cup S) \geq |S| + 2.$$

Consequently

$$\begin{aligned} |S_0 \cup \{c\} \cup S| &= |S_0| + 1 + |S| \geq q(G - S_0 \cup \{c\} \cup S) \\ &= q(G - S_0) - 1 + q(C_i - \{c\} \cup S) \geq |S_0| + 1 + |S|, \end{aligned}$$

so in (*) we have equality for the set $S_0 \cup \{c\} \cup S$ as well. This contradicts the maximality of S_0.

(iii) *G contains m independent edges of the form $s_i c_i$, $s_i \in S_0$ and $c_i \in C_i$, $i = 1, 2, \ldots, m$.* To show this let us consider the bipartite graph $H = G_2(n, m)$ with vertex classes $\{C_1, C_2, \ldots, C_m\}$ and $V_2 = S_0$, in which C_i is joined to a vertex $s \in S_0$ if and only if G contains an edge from s to C_i. The assertion above is true iff H has a 1-factor, that is a matching from V_1 to V_2. Fortunately we have got the weapon to check this: Hall's theorem. Given $A \subset V_1$, put $B = \Gamma_H(A) \subset V_2$ (see Figure III.7). Then (*) implies that

$$|A| \leq q(G - B) \leq |B|.$$

Hence the graph H satisfies Hall's condition, so it has 1-factor.

We are almost home. To complete the proof we just put together the information from (i), (ii) and (iii). We start with the m independent edges $s_i c_i$, $s_i \in S_0$, $c_i \in C_i$. Adding to this set of edges a 1-factor of each $C_i - c_i$, $1 \leq i \leq m$, and a 1-factor of each D_j, $1 \leq j \leq k$, we arrive at a 1-factor of G. □

It is once again very easy to obtain a defect form of the above result.

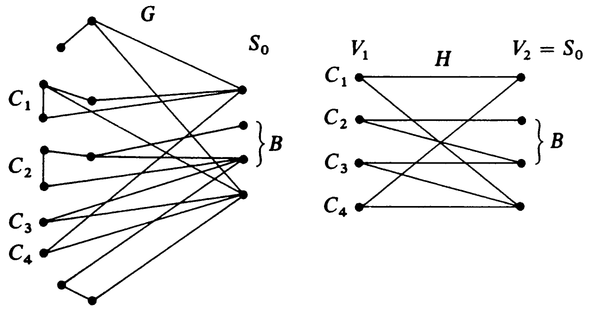


Figure III.7. The construction of H from G. The set $A = \{C_2, C_3\}$ determines $B \subset S_0$ by the rule $B = \Gamma_H(A)$.

Corollary 13. *A graph G contains a set of independent edges covering all but at most d of the vertices iff*

$$q(G - S) \leq |S| + d \quad \textit{for every } S \subset V(G).$$

PROOF. Since the number of vertices not covered by a set of independent edges is congruent to $|G|$ modulo 2, we may assume that

$$d \equiv |G| \pmod 2.$$

Put $H = G + K^d$, that is let H be obtained from G by adding to it a set W of d vertices and joining every new vertex to every other vertex, old and new. Then G contains a set of independent edges covering all but d of the vertices iff H has got a 1-factor. When does (*) hold for H? If $W - S' \neq \varnothing$ then $H - S'$ is connected so $q(H - S') \leq 1$ and then (*) does hold; if $W \subset S'$ then with $S = S' - W$ we have $q(H - S') = q(G - \{S' \backslash W\}) = q(G - S)$, so (*) is equivalent to

$$q(G - S) \leq |S'| = |S| + d. \qquad \square$$

EXERCISES

1.⁻ Suppose F is a set of edges after whose deletion there is no flow from s to t with positive value. Prove that F contains a cut separating s from t.

2.⁻ By summing an appropriate set of equations show that the capacity of a cut is at least as large as the maximum of the flow value.

3.⁻ Let $\vec{G} = (V, \vec{E})$ be a directed graph and let c be an extended-real-valued capacity function on $\vec{E}$. (Thus $c(x, y)$ is a non-negative real or $+\infty$.) Let s and t be two vertices. Prove that either there is a flow from s to t with infinite value or else there is a flow with maximal finite value.

4. By successively reducing the number of circular flows in $\vec{G}$, prove that there is a maximal flow without circular flows in which no current enters the source and no current leaves the sink.

5. Use the method of Exercise 4 to show that if the capacity function is integral then there is a maximal flow that is also integral.

6. Formulate and prove the max-flow min-cut theorem for multiple sources and sinks with bounds on the capacity of the edges and vertices.

7. (Circulation theorem). A *circulation* in a directed graph $\vec{G}$ is a flow without a source and a sink. Given a lower capacity $l(x, y)$ and an upper capacity $c(x, y)$ for each edge $\overrightarrow{xy}$ with $0 \leq l(x, y) \leq c(x, y)$ we call a circulation g *feasible* if

$$l(x, y) \leq g(x, y) \leq c(x, y)$$

for every edge $\overrightarrow{xy}$. Prove that there exists a feasible circulation iff

$$l(S, \bar{S}) \leq c(\bar{S}, S) \quad \text{for every } S \subset V.$$

[Note that the necessity of the condition is trivial since in a feasible circulation the function l forces at least $l(S, \bar{S})$ current from S to $\bar{S}$ and the function c allows at most $c(\bar{S}, S)$ current into $\bar{S}$ from S. To prove the sufficiency, adjoin a sink s and a source t to $\vec{G}$, send an edge from s to each vertex of $\vec{G}$ and send an edge from each vertex of $\vec{G}$ to t. Define a capacity function c^* on the edges of the new graph $\vec{G}^*$ by putting $c^*(x, y) = c(x, y) - l(x, y)$, $c^*(s, x) = l(V, x)$ and $c^*(x, t) = l(x, V)$. Then the relation

$$f(x, y) = g(x, y) - l(x, y)$$

sets up a 1-1 correspondence between the feasible circulations g in $\vec{G}$ and flows f in $\vec{G}^*$ from s to t with value $l(V, V)$. Rewrite the condition given in the max-flow min-cut theorem in the form required in this result.]

8.[+] Let H be a bipartite multigraph without loops, with vertex classes V_1 and V_2. (Thus H may contain multiple edges, that is two vertices belonging to different classes may be joined by several edges, which are said to be *parallel*.) As usual, given a vertex x we denote by $\Gamma(x)$ the set of edges incident with x and by $d(x) = |\Gamma(x)|$ the degree of x. Prove that, given any natural number k, the set E of edges can be partitioned into sets $E_1, E_2, \ldots, E_k$ such that for every vertex x and every set E_i we have

$$\left\lfloor \frac{d(x)}{k} \right\rfloor \leq |\Gamma(x) \cap E_i| \leq \left\lceil \frac{d(x)}{k} \right\rceil,$$

where, as in the rest of the book, $\lceil z \rceil$ is the least integer not less than z and $\lfloor z \rfloor = -\lceil -z \rceil$.

Thus if we think of the partition $\bigcup_1^k E_i$ as a colouring of the edges with k colours, then the colouring is *equitable* in the sense that in each vertex the distribution of colours is as equal as possible. [*Hint*. Construct a directed graph $\vec{H} = (V_1 \cup V_2, \vec{E})$ from H by sending an edge from x to y iff $x \in V_1$, $y \in V_2$ and H contains at least one

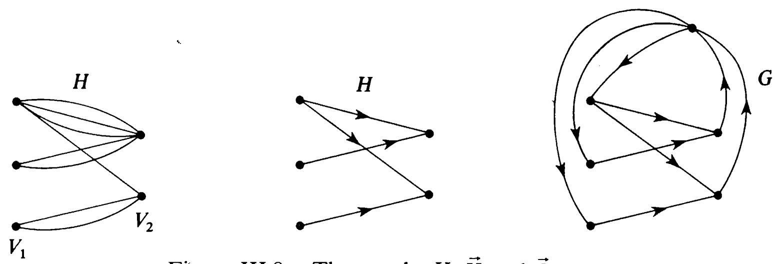


Figure III.8. The graphs H, $\vec{H}$ and $\vec{G}$.

xy edge. Let $\vec{G}$ be obtained from $\vec{H}$ by adding a vertex u and all the edges $\vec{ux}$, $\vec{yu}$, for $x \in V_1$ and $y \in V_2$, as shown in Figure III.8. Define an appropriate upper and lower capacity for each edge of $\vec{G}$ and prove that there is a feasible integral circulation. Use this circulation to define one of the colour classes.]

9. (Exercise 8^+ continued) Show that we may require that, in addition to the property above, the colour classes are as equal as possible, say $|E_1| \leq |E_2| \leq \cdots \leq |E_k| \leq |E_1| + 1$, and in each set of parallel edges the distribution of colours is as equal as possible.

10. Let $d_1 \leq d_2 \leq \cdots \leq d_n$ be the degree sequence of a graph G. Suppose that

$$d_j \geq j + k - 1 \quad \text{for } j = 1, 2, \ldots, n - 1 - d_{n-k+1}.$$

11. Prove that G is k-connected.

Let k and l be integers $1 \leq k \leq l$. Construct graphs G_1, G_2 and G_3 such that
(i) $\kappa(G_1) = k$ and $\lambda(G_1) = l$,
(ii) $\kappa(G_2) = k$ and $\kappa(G_2 - x) = l$ for some vertex x,
(iii) $\kappa(G_3 - x) = k$ and $\lambda(G_3 - xy) = l$ for some edge xy.

12. Given $U \subset V(G)$ and a vertex $x \in V(G) - U$, an *x-U fan* is a set of $|U|$ paths from x to U, any two of which have exactly the vertex x in common. Prove that a graph G is k-connected iff $|G| \geq k + 1$ and for any k-set $U \subset V(G)$ and vertex x not in U, there is an x-U fan in G. [*Hint*. Given a pair (x, U), add a vertex u to G and join it to each vertex in U. Check that the new graph is k-connected if G is. Apply Menger for x and u.]

13. Prove that if G is k-connected ($k \geq 2$) and $\{x_1, x_2, \ldots, x_k\} \subset V(G)$ then there is a cycle in G that contains all x_i, $1 \leq i \leq k$.

14. The *line graph* $L(G)$ of a graph $G = (V, E)$ has vertex set E and two vertices $\alpha, \beta \in E$ are adjacent iff they have exactly one vertex of G in common. By applying the vertex form of Menger's theorem to the line graph $L(G)$, prove that the vertex form of Menger's theorem implies the edge form.

15. Show that if $\lambda(G) = k \geq 2$ then the deletion of k edges from G results in at most 2 components. Is there a similar result for vertex-connectivity?

16. Let G be a connected graph with minimum degree $\delta(G) = k \geq 1$. Prove that G contains a path $x_1x_2\ldots x_k$ such that $G - \{x_1, x_2, \ldots, x_k\}$ is also connected. [*Hint*. Let $x_1x_2\ldots x_l$ be a longest path. Note that $l \geq k + 1$. Suppose that $G - \{x_1, x_2, \ldots, x_k\}$ is disconnected and let $y_0y_1\ldots y_m$ be a longest path in a component C not containing $x_{k+1}x_{k+2}\ldots x_l$. Then $d_C(y_0) \leq m$ but y_0 cannot be joined to $k - m$ of the vertices $x_1, \ldots, x_k$.]

17. Let $G = G_2(m, n)$ be a bipartite graph with vertex classes V_1 and V_2 containing a matching from V_1 to V_2.
(i) Prove that there is a vertex $x \in V_1$ such that for every edge xy there is a matching from V_1 to V_2 that contains xy.
(ii) Deduce that if $d(x) = d$ for every $x \in V_1$ then G contains at least $d!$ matchings if $d \leq m$ and at least $d(d-1)\cdots(d-m+1)$ matchings if $d > m$.

18. Let $A = (a_{ij})_1^n$ be an $n \times n$ *stochastic matrix*, that is let $a_{ij} \geq 0$ and $\sum_{i=1}^n a_{ij} = \sum_{j=1}^n a_{ij} = 1$ for all i, j. Show that A is in the convex hull of $n \times n$ permutation

matrices, i.e. there are $\lambda_i \geq 0, \sum_1^m \lambda_i = 1$, and permutation matrices $P_1, P_2, \ldots, P_m$ such that $A = \sum_1^m \lambda_i P_i$. [Let $a_{ij}^* = \lceil a_{ij} \rceil$, $A^* = (a_{ij}^*)_1^n$, and let $G = G_2(n, n)$ be the bipartite graph naturally associated with A^*. Show that G has a complete matching and deduce that there are a permutation matrix P and a real $\lambda, 0 < \lambda \leq 1$, such that $A - \lambda P = B = (b_{ij})_1^n$ is such that $b_{ij} \geq 0, \sum_{i=1}^n b_{ij} = \sum_{j=1}^n b_{ij} = 1 - \lambda$ for all i, j, *and* B has at least one more 0 entry than A.]

19. An $r \times s$ *Latin rectangle* based on $1, 2, \ldots, n$ is an $r \times s$ matrix $A = (a_{ij})$ such that each entry is one of the integers $1, 2, \ldots, n$ and each integer occurs in each row and column at most once. Prove that every $r \times n$ Latin rectangle A can be extended to an $n \times n$ *Latin square.* [*Hint.* Assume that $r < n$ and extend A to an $(r + 1) \times n$ Latin rectangle. Let A_j be the set of possible values of $a_{r+1,j}$, that is let $A_j = \{k: 1 \leq k \leq n,\ k \neq a_{ij}\}$. Check that $\{A_j: 1 \leq j \leq n\}$ has a set of distinct representatives.]

    Prove that there are at least $n!\,(n - 1)! \ldots (n - r + 1)!$ distinct $r + n$ Latin rectangles based on $1, 2, \ldots, n$. [*Hint.* Apply Exercise 17(ii).]

20. Let A be an $r \times s$ Latin rectangle and denote by $A(i)$ the number of times the symbol i occurs in A. Show that A can be extended to an $n \times n$ Latin square iff $A(i) \geq r + s - n$ for every $i = 1, 2, \ldots, n$.

21. Prove the following form of the Schröder–Bernstein theorem. Let G be a bipartite graph with vertex classes X and Y having arbitrary cardinalities. Let $A \subset X$ and $B \subset Y$. Suppose there are complete matchings from A into Y and from B into X. Prove that G contains a set of independent edges covering all the vertices of $A \cup B$. [*Hint.* Consider the components of the union of the matchings.]

22. Deduce from Exercise 20 that every bipartite graph contains a set of independent edges such that each vertex of maximum degree (that is degree $\Delta(G)$) is incident with one of the edges. Deduce that a non-empty regular bipartite graph has a 1-factor.

23. We say that G is an $(r, r - k)$*-regular* graph if $r - k \leq \delta(G) \leq \Delta(G) \leq r$. Prove that for $1 \leq k \leq s \leq r$ every $(r, r - k)$-regular graph contains an $(s, s - k)$-regular factor. [*Hint.* Assume $s = r - 1$. Take a minimal $(r, r - k)$-regular factor. Note that in this factor no two vertices of degree r are adjacent. Remove a set of independent edges covering the vertices of degree r.]

24.$^+$ Recall Tychonov's theorem from general topology: the product of an arbitrary family of compact spaces is compact. Deduce from this the following extension of Hall's theorem.

    Let G be an infinite bipartite graph with vertex classes X and Y, such that each vertex in X is incident with finitely many edges. Then there is a complete matching from X into Y iff $|\Gamma(A)| \geq |A|$ for every finite subset A of X.

    Show that the finiteness condition cannot be omitted.

25. Prove that a 2-edge-connected cubic graph has a 1-factor. [This result is called Petersen's theorem. In order to prove it check that the condition of Tutte's theorem is satisfied. If $\varnothing \neq S \subset V(G)$ and C is an odd component of $G - S$ then there are at least two $S - C$ edges, since G is 2-edge-connected. Furthermore, since G is cubic, there are at least three $S - C$ edges. Deduce that $q(G - S) \leq |S|$.]

26. Imitate the second proof of Theorem 7 to give a direct proof of the case $d = 0$ of Theorem 10 and then deduce from it the general case $d \geq 0$.

27. Let G be a graph of order n with at most $r \geq 2$ independent vertices. Prove that if $\vec{G}$ is any orientation of G which does not contain a directed cycle (*acyclic* orientation) then $\vec{G}$ contains a directed path of length at least $\lceil n/r \rceil - 1$. [*Hint.* Apply Dilworth's theorem, Theorem 11.]

28. Deduce from Exercise 27 the following result. Given a set of $rk + 1$ distinct natural numbers, either there exists a set of $r + 1$ numbers, none of which divides any of the other r numbers, or else there exists a sequence $a_0 < a_1 < \cdots < a_k$ such that if $0 \leq i < j \leq k$ then a_i divides a_j.

29. Describe all maximal graphs of order $n = 2l$ which do not contain a 1-factor. [*Hint.* Read it out of Tutte's theorem (Theorem 12).]

30. (cf. Exercise 29) Describe all maximal graphs of order n which contain at most k independent edges. [*Hint.* Apply Corollary 13.]

31. Make use of Exercise 30 and the convexity of the binomial coefficient $\binom{x}{2}$, $x \geq 2$, to prove that if $n \geq k + 1$ then the maximal size of a graph of order n with at most k independent edges is

$$\max\left\{\binom{2k+1}{2}, \binom{k}{2} + k(n-k)\right\}.$$

Show also that the extremal graphs (that is the graphs for which equality holds) are one or both of the graphs $K^{2k+1} \cup E^{n-2k-1}$ and $K^k + E^{n-k}$. (See Figure III.9.)

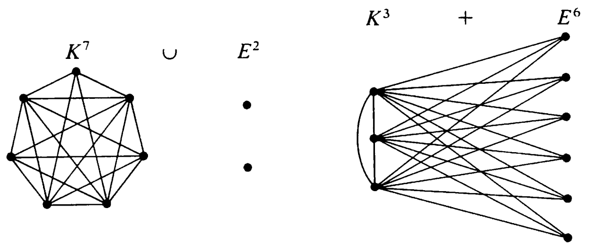


Figure III.9. For $k = 3$, $n = 9$ there are two extremal graphs: $K^7 \cup E^2$ and $K^3 + E^6$.

32. Call a sequence $d_1, d_2, \ldots, d_n$ of integers *graphic* if there is a graph G with vertex set $V(G) = \{x_1, x_2, \ldots, x_n\}$ such that $d(x_i) = d_i$, $1 \leq i \leq n$. (The graph G is said to *realize* $(d_i)_1^n$.) Show that $d_1 \geq d_2 \geq \cdots \geq d_n$ is graphic iff so is the sequence

$$d_2 - 1, d_3 - 1, \ldots, d_{d_1+1} - 1, d_{d_1+2}, d_{d_1+3}, \ldots, d_n.$$

33. Use the algorithm given in the previous exercise to decide which of the following sequences are graphic: 5, 4, 3, 2, 2, 2; 5, 4, 4, 2, 2, 1; 4, 4, 3, 3, 2, 2, 2 and 5, 5, 5, 4, 2, 1, 1, 1. Draw the graphs realizing the appropriate sequences constructed by the algorithm.

Notes

The basic book on flows is L. R. Ford, Jnr. and D. R. Fulkerson, *Flows in Networks*, Princeton University Press, Princeton, 1962. It not only contains all the results mentioned in the chapter concerning flows and circulations, but also a number of applications to standard optimization problems.

The fundamental theorems of Menger, Hall and Tutte are in K. Menger, Zur allgemeinen Kurventheorie, *Fund. Math.* **10** (1927) 96–115, P. Hall, On representatives of subsets, *J. London Math. Soc.* **10** (1935) 26–30), and W. T. Tutte, A factorization of linear graphs, *J. London Math. Soc.* **22** (1947) 107–111. The proof of Tutte's theorem we gave is due to T. Gallai, Neuer Beweis eines Tutte'schen Satzes, *Magyar Tud. Akad. Közl.* **8** (1963) 135–139, and was rediscovered independently by I. Anderson, Perfect matchings in a graph, *J. Combinatorial Theory* Ser. B **10** (1971) 183–186 and W. Mader, Grad und lokaler Zusammenhang in endlichen Graphen, *Math. Ann.* **205** (1973) 9–11. The result in Exercises 8 and 9 is due to D. de Werra, Multigraphs with quasiweak odd cycles, *J. Combinatorial Theory* Ser. B **23** (1977) 75–82.

A slightly simpler form of the result in Exercise 23 is due to W. T. Tutte; the proof indicated in the hint was found by C. Thomassen. An extensive survey of results concerning connectivity and matching can be found in Chapters I and II of B. Bollobás, *Extremal Graph Theory*, Academic Press, London and New York, 1978.

CHAPTER IV

Extremal Problems

At least how many edges are in a graph of order n if it is forced to contain a path of length l? A cycle of length at least l? A cycle of length at most l? A complete graph K^r? These questions are special cases of the so called *forbidden subgraph problem*: given a graph F, determine $\mathrm{ex}(n; F)$, the maximal number of edges in a graph of order n not containing F. The forbidden subgraph problem is a prime example of the rather large family of *extremal problems* in graph theory. In this chapter we present the fundamental result concerning the forbidden subgraph problem and discuss some closely related extremal questions. We have to emphasize that we do not intend to provide an overall view of extremal graph theory.

Before going into the details, it is appropriate to say a few words about the terminology. If for a given class of graphs a certain graph parameter, say the number of edges or the minimum degree, is at most some number f, then the graphs for which equality holds are the *extremal graphs* of the inequality. As a trivial example, note that an acyclic graph of order n has at most $n - 1$ edges, and the extremal graphs are the trees of order n. In the forbidden subgraph problem a graph is *extremal* if it does not contain F and has $\mathrm{ex}(n; F)$ edges. In the previous chapter Tutte's factor theorem enabled us to solve a beautiful extremal problem: how many edges guarantee $k + 1$ independent edges? In this case F consists of $k + 1$ independent edges, that is $F = (k + 1)K^2$. Exercise 31 of Chapter III claims that for $n \geq 2k + 1$ the extremal graphs of $\mathrm{ex}(n; F)$ are $K^{2k+1} \cup E^{n-2k-1}$ and/or $K^k + E^{n-k}$.

The material in this chapter falls conveniently into two parts: the odd sections concern paths and cycles while the even ones are about complete subgraphs. We have chosen to alternate the topics in order to have the simpler results first, as in most other chapters.

The first section is about paths and cycles (short and long) in graphs of

large size. Among others we shall give a good bound on $\mathrm{ex}(n; P^l)$, the maximal number of edges in a graph of order n without a path of length l. We shall also present some fundamental results about Hamilton cycles.

One of the best known theorems in graph theory, Turán's theorem, determines the function $\mathrm{ex}(n; K^r)$. The second section is devoted to this theorem together with some related results.

When discussing $\mathrm{ex}(n; P^l)$ and $\mathrm{ex}(n; K^r)$, we mostly care about the case when n is large compared to l and r. We get rather different problems if F and G have the same order. A prime example of these problems will be discussed in the third section, the problem of Hamilton cycles. Over the years considerable effort has gone into the solution of this problem and in a certain rather narrow sense the present answers are satisfactory.

The fourth section is devoted to a deep and surprising theorem of Erdös and Stone. The theorem concerns $\mathrm{ex}(n; F)$, where F is a complete r-partite graph with t vertices in each class, but as an immediate corollary of this result one can determine $\lim_{n\to\infty} \mathrm{ex}(n; F)/n^2$ for every graph F.

§1 Paths and Cycles

In its natural formulation our first theorem gives a lower bound on the order of a graph in terms of the minimum degree and the *girth*, the length of a shortest cycle. Equivalently, the result gives an upper bound on the girth in terms of the order and minimum degree.

Theorem 1. *For $g \geq 3$ and $\delta \geq 3$ put*

$$n_0(g, \delta) = \begin{cases} 1 + \dfrac{\delta}{\delta - 2}\{(\delta - 1)^{(g-1)/2} - 1\} & \text{if } g \text{ is odd,} \\ \dfrac{2}{\delta - 2}\{(\delta - 1)^{g/2} - 1\} & \text{if } g \text{ is even.} \end{cases}$$

Then a graph G with minimum degree δ and girth g has at least $n_0(g, \delta)$ vertices.

Proof. Suppose first that g is odd, say $g = 2d + 1$, $d \geq 1$. Pick a vertex x. There is no vertex z for which G contains two distinct z-x paths of length at most d, since otherwise G has a cycle of length at most $2d$. Consequently there are at least δ vertices at distance 1 from x, at least $\delta(\delta - 1)$ vertices at distance 2, and so on, at least $\delta(\delta - 1)^{d-1}$ vertices at distance d from x (Figure IV.1). Thus

$$n \geq 1 + \delta + \delta(\delta - 1) + \cdots + \delta(\delta - 1)^{d-1},$$

as claimed.



Figure IV.1. The cases $\delta = g = 5$ and $\delta = 4, g = 6$.

Suppose now that g is even, say $g = 2d$. Pick two adjacent vertices, say x and y. Then there are $2(\delta - 1)$ vertices at distance 1 from $\{x, y\}$, $2(\delta - 1)^2$ vertices at distance 2, and so on, $2(\delta - 1)^{d-1}$ vertices at distance $d - 1$ from $\{x, y\}$, implying the required inequality. □

Let G_0 be an *extremal graph* of Theorem 3, that is a graph with parameters δ and g, for which equality holds. The proof above implies that G_0 is *regular of degree* δ; if $g = 2d + 1$ then G_0 has *diameter* d and if $g = 2d$ then every vertex is within distance $d - 1$ of each pair of adjacent vertices. It is easily seen that $n_0(g, \delta)$ is also the *maximal* number for which there is a graph with *maximum degree* δ having the latter property (Exercise 1). We call G_0 a *Moore graph of degree* δ *and girth* g, or, if $g = 2d + 1$, a *Moore graph of degree* δ *and diameter* d. In Chapter VIII we shall use algebraic methods to investigate Moore graphs.

Let us see now what we can say about *long* cycles and paths in a graph. If a graph of order n is Hamiltonian then its *circumference*, that is the length of a longest cycle, is n, while the length of a longest path is $n - 1$. However, every non-Hamiltonian *connected* graph contains at least as long paths as the circumference of the graph. Indeed, if $C = x_1x_2 \dots x_l$ is a longest cycle and $l < n$ then there is a vertex y not on C adjacent to a vertex of C, say x_1, and then $yx_1x_2 \dots x_l$ is a path of length l.

Theorem 2. *Let G be a connected graph of order $n \geq 3$ such that for any two non-adjacent vertices x and y we have*

$$d(x) + d(y) \geq k.$$

If $k = n$ then G is Hamiltonian and if $k < n$ then G contains a path of length k. and a cycle of length at least $(k + 2)/2$.

PROOF. Assume that G is not Hamiltonian and let $P = x_1x_2 \dots x_l$ be a longest path in G. The maximality of P implies that the neighbours of x_1 and x_l are vertices of P. As G does not contain a cycle of length l, x_1 is not adjacent to x_l. Furthermore the path P cannot contain vertices x_i and x_{i+1} such that x_1 is adjacent to x_{i+1} and x_l is adjacent to x_i since otherwise $x_1x_2 \dots x_ix_nx_{n-1} \dots x_{i+1}$ is a cycle of length l (Figure IV.2). Consequently the sets

$$\Gamma(x_1) = \{x_j : x_1x_j \in E(G)\} \quad \text{and} \quad \Gamma^+(x_l) = \{x_{i+1} : x_ix_l \in E(G)\}$$

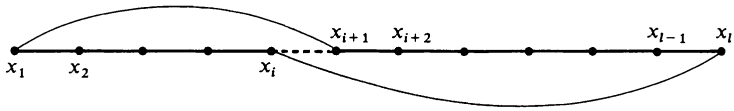


Figure IV.2. The construction of a cycle of length l.

are disjoint subsets of $\{x_2, x_3, \ldots, x_l\}$ and so

$$k \leq d(x_1) + d(x_l) \leq l - 1 \leq n - 1.$$

The first two assertions of the theorem follow from this inequality. If $k = n$, this is impossible, so G is Hamiltonian. If $k < n$ then P has length $l - 1 \geq k$.

Finally, the assertion about cycles is even simpler. Assume that $d(x_1) \geq d(x_l)$ so $d(x_1) \geq \lceil k/2 \rceil$, where $\lceil z \rceil$ denotes the least integer not less than z. Put $t = \max\{i : x_1 x_i \in E(G)\}$. Then $t \geq d(x_1) + 1 \geq \lceil k/2 \rceil + 1$ and G contains the cycle $x_1 x_2 \ldots x_t$ of length t. □

In §3 we shall make use of the proof of this theorem to obtain detailed information about graphs without long cycles and paths. For the moment we confine ourselves to noting two consequences of Theorem 2.

Theorem 3. *Let G be a graph of order n without a path of length $k(\geq 1)$. Then*

$$e(G) \leq \frac{k-1}{2} n.$$

A graph is an extremal graph (that is equality holds for it) iff all its components are complete graphs of order k.

Proof. We fix k and apply induction on n. The assertion is clearly true if $n \leq k$. Assume now that $n > k$ and the assertion holds for smaller values of n.

If G is disconnected, the induction hypothesis implies the result. Now if G is connected, it contains no K^k and, by Theorem 1, it has a vertex x of degree at most $(k - 1)/2$. Since $G - x$ is not an extremal graph

$$e(G) \leq d(x) + e(G - x) < \frac{k-1}{2} + \frac{k-1}{2}(n-1) = \frac{k-1}{2} n. \qquad \square$$

Theorem 4. *Let G be a graph of order n in which every cycle has length at most k $(k \geq 2)$. Then*

$$e(G) \leq \frac{k}{2}(n-1).$$

A graph is an extremal graph iff it is connected and all its blocks are complete graphs of order k. □

The proof of this result is somewhat more involved than that of Theorem 3. Since a convenient way of presenting it uses 'simple transforms' to be introduced in §3, the proof is left as an exercise (Exercise 25) with a detailed hint.

§2 Complete Subgraphs

What is $\mathrm{ex}(n; K^r)$, the maximal number of edges in a graph of order n not containing a K^r, a complete graph of order r? If G is $(r-1)$-partite then it does not contain a K^r since every vertex class of G contains at most one vertex of a complete subgraph. Thus $\mathrm{ex}(n; K^r)$ is at least as large as the maximal size of an $(r-1)$-partite graph of order n. In fact, there is a unique $(r-1)$-partite graph of order n that has maximal size. This graph is the complete $(r-1)$-partite graph $T_{r-1}(n)$ that has $\lfloor (n+k-1)/(r-1) \rfloor$ vertices in the kth class (Figure IV.3). In other words, $T_{r-1}(n)$ is the complete $(r-1)$-partite graph of order n whose classes are as equal as possible: say there are n_k vertices in the kth class and $n_1 \le n_2 \le \ldots \le n_{r-1} \le n_1 + 1$. Indeed, let G be an $(r-1)$-partite graph of order n and maximal size.

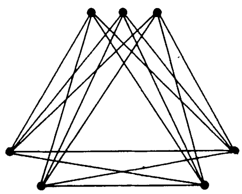


Figure IV.3. The Turán graph $T_3(7)$.

Clearly G is a complete $(r-1)$-partite graph. Suppose the classes are not as equal as possible, say there are m_1 vertices in the first class and $m_2 \ge m_1 + 2$ in the second. This is impossible since by transferring one vertex from the second class to the first we would increase the number of edges by $(m_1+1)(m_2-1) - m_1 m_2 = m_2 - m_1 - 1 \ge 1$.

The number of edges in $T_{r-1}(n)$ is usually denoted by $t_{r-1}(n)$; thus $t_2(n) = \lfloor n^2/4 \rfloor$. A fundamental theorem of Turán states that the trivial inequality $\mathrm{ex}(n; K^r) \ge t_{r-1}(n)$ is, in fact, an equality for every n and r. We shall show first that the degree sequence of a graph without a K^r is dominated by the degree sequence of an $(r-1)$-partite graph. In view of the remarks above this will imply Turán's theorem.

Theorem 5. *Let G be a graph with vertex set V that does not contain K^r, a complete graph of order r. Then there is an $(r-1)$-partite graph H with vertex set V such that for every vertex $z \in V$ we have*

$$d_G(z) \le d_H(z).$$

If G is not a complete $(r-1)$-partite graph then there is at least one vertex z for which the inequality above is strict.

PROOF. We shall apply induction on r. For $r = 2$ there is nothing to prove since G is the empty graph E^n, which is 1-partite. Assume now that $r \geq 3$ and the assertion holds for smaller values of r.

Pick a vertex $x \in V$ for which $d_G(x)$ is maximal and denote by W the set of vertices of G that are joined to x. Then $G_0 = G[W]$ does not contain a K^{r-1} otherwise with x it would form a K^r. By the induction hypothesis we can replace G_0 by an $(r-2)$-partite graph H_0 with vertex set W in such a way that $d_{G_0}(y) \leq d_{H_0}(y)$ for every $y \in W$. Add to H_0 the vertices in $V - W$ and join each vertex in $V - W$ to each vertex in W. To complete the proof let us check that the graph H obtained in this way has the required properties.

If $z \in V - W$ then $d_H(z) = d_H(x) = d_G(x) \geq d_G(z)$ and if $z \in W$ then $d_H(z) = d_{H_0}(z) + n - |W| \geq d_{G_0}(z) + n - |W| \geq d_G(z)$. Thus $d_G(z) \leq d_H(z)$ holds for every $z \in V$.

Can H contain a complete subgraph of order r? Clearly not, since a complete subgraph of H has at most one vertex in $V - W$ and most $r - 2$ vertices in W.

What can we say about G is $e(G) = e(H)$? Then $e(G_0) = e(H_0)$, so by the induction hypothesis G_0 is a complete $(r-2)$-partite graph and

$$e_G(W, \overline{W}) = e_H(W, \overline{W}) = |W||\overline{W}|.$$

Hence

$$e(G[\overline{W}]) = e(H[\overline{W}]) = 0,$$

implying that G is a complete $(r-1)$-partite graph. □

We now give *Turán's theorem* in its original form.

Theorem 6. $\text{ex}(n; K^r) = t_{r-1}(n)$ *and* $T_{r-1}(n)$ *is the unique graph of order n and size $t_{r-1}(n)$ that does not contain a complete graph of order r.*

1ST PROOF. Since $T_{r-1}(n)$ is the unique $(r-1)$-partite graph of order n and maximum size, both assertions follow from Theorem 5. □

In order to emphasize the significance of Theorem 6, we prove it from first principles also. This proof is based on the fact that $T_{r-1}(n)$ is almost regular: $\delta(T_{r-1}(n)) = n - \lceil n/(r-1) \rceil$, $\Delta(T_{r-1}(n)) = n - \lfloor n/(r-1) \rfloor$, and if x is a vertex of minimum degree in $T_{r-1}(n)$ then $T_{r-1}(n) - x$ is exactly $T_{r-1}(n-1)$.

2ND PROOF. Fix $r \geq 3$ and apply induction on n. For $n \leq r$ the assertion is trivial so suppose that $n > r$ and the theorem holds for smaller values of n.

Suppose G has n vertices, $t_{r-1}(n)$ edges and it contains no K^r. As $T_{r-1}(n)$ is a maximal graph without a K^r (that is no edge can be added to it without

creating a K^r) the induction step will follow if we show that G is exactly a $T_{r-1}(n)$. Since the degrees in $T_{r-1}(n)$ differ by at most 1, we have

$$\delta(G) \leq \delta(T_{r-1}(n)) \leq \Delta(T_{r-1}(n)) \leq \Delta(G).$$

Let x be a vertex of G with degree $d(x) = \delta(G) \leq \delta(T_{r-1}(n))$. Then

$$e(G - x) = e(G) - d(x) \geq e(T_{r-1}(n - 1)),$$

so by the induction hypothesis $G_x = G - x$ is exactly a $T_{r-1}(n - 1)$.

A smallest vertex class of G_x contains $\lfloor (n - 1)/(r - 1) \rfloor$ vertices and the vertex x is joined to all but

$$n - 1 - \left(n - \left\lceil \frac{n}{r-1} \right\rceil\right) = \left\lfloor \frac{n-1}{r-1} \right\rfloor$$

vertices of G_x. Since x cannot be joined to a vertex in each class of G_x, it has to be joined to all vertices of G_x save the vertices in a smallest vertex class. This shows that G is exactly a $T_{r-1}(n)$ graph. □

The proof above can easily be adapted to give a number of related results (cf. Exercises 7, 8).

Now let us turn to the *problem of Zarankiewicz*, which is the analog of Turán's problem in bipartite graphs. Write $G_2(m, n)$ for a bipartite graph with m vertices in the first class and n in the second. What is the maximum size of a $G_2(m, n)$ if it does not contain a complete bipartite graph with s vertices in the first class and t in the second? This maximum is usually denoted by $z(m, n; s, t)$. The following simple lemma seems to imply a very good upper bound for the function $z(m, n; s, t)$.

Lemma 7. *Let m, n, s, t, r, k be integers, $2 \leq s \leq m$, $2 \leq t \leq n$, $0 \leq k$, $0 \leq r < m$, and let $G = G_2(m, n)$ be a graph of size $z = my = km + r$ without a $K(s, t)$ subgraph having s vertices in the first class and t in the second. Then*

$$m\binom{y}{t} \leq (m - r)\binom{k}{t} + r\binom{k+1}{t} \leq (s - 1)\binom{n}{t}. \tag{1}$$

Proof. Denote by V_1 and V_2 the vertex classes of G. We shall say that a t-set (i.e., a set with t elements) T of V_2 *belongs* to a vertex $x \in V_1$ if x is joined to every vertex in T. The number of t-sets belonging to a vertex $x \in V_1$ is $\binom{d(x)}{t}$. Since the assumption on G is exactly that each t-set in V_2 belongs to at most $s - 1$ vertices of V_1, we find that

$$\sum_{x \in V_1} \binom{d(x)}{t} \leq (s - 1)\binom{n}{t}. \tag{2}$$

As $\sum_{x \in V_1} d(x) = z = my = km + r$, $0 \leq r < m$ and $f(u) = \binom{u}{t}$ is a convex function of u for $u \geq t$, inequality (2) implies (1). □

Theorem 8. $z(m, n; s, t) \leq (s - 1)^{1/t}(n - t + 1)m^{1-(1/t)} + (t - 1)m.$

PROOF. Let $G = G_2(m, n)$ be an extremal graph for the function $z(m, n; s, t)$, that is let G be a bipartite graph of size $z(m, n; s, t) = my$ without a $K(s, t)$ subgraph. As $y \leq n$, inequality (1) implies

$$(y - (t - 1))^t \leq (s - 1)(n - (t - 1))^t m^{-1}. \qquad \square$$

As a consequence of Theorem 8 we see that if $t \geq 2$ and $c > (t - 1)^{1/t}$ are fixed then for every sufficiently large n we have

$$z(n, n; t, t) < cn^{2-(1/t)}. \tag{3}$$

The method of Lemma 7 gives also an upper bound for $\mathrm{ex}(n; K_2(t))$, the maximum number of edges in a graph of order n without a complete t by t bipartite subgraph.

Theorem 9. $\mathrm{ex}(n; K_2(t)) \leq \frac{1}{2}(t - 1)^{1/t}(n - t + 1)n^{1-(1/t)} + \frac{1}{2}(t - 1)n$
$< \frac{1}{2}(t - 1)^{1/t}n^{2-(1/t)} + \frac{1}{2}tn.$

PROOF. Let G be an extremal graph. As in Lemma 7, let us say that a t-set of the vertices belongs to a vertex x if x is joined to every vertex of the t-set. Since G does not contain a $K_2(t)$, every t-set belongs to at most $t - 1$ vertices. Therefore if G has degree sequence $(d_i)_1^n$ then $\sum_1^n d_i = 2\,\mathrm{ex}(n; K_2(t))$ and

$$\sum_1^n \binom{d_i}{t} \leq (t - 1)\binom{n}{t},$$

implying the result. $\square$

It is very likely that inequality (3) gives the correct order of magnitude of $z(n, n; t, t)$ for every $t \geq 2$ as $n \to \infty$ but this has been proved only for $t = 2$ and 3. In fact, it is rather hard to find non-trivial lower bounds for $z(m, n; s, t)$. In Chapter VII we shall use the probabilistic method to obtain a lower bound. To conclude the section we prove an elegant result for $t = 2$ that indicates the connection between the problem of Zarankiewicz and designs, in particular, projective spaces.

Theorem 10. $z(n, n; 2, 2) \leq \frac{1}{2}n\{1 + (4n - 3)^{1/2}\}$ *and equality holds for infinitely many values of n.*

PROOF. Note that with $z = \frac{1}{2}n\{1 + (4n - 3)^{1/2}\}$ we have

$$(z - n)z = n^2(n - 1).$$

Suppose that there is a bipartite graph $G = G_2(n, n)$ of size greater than z that does not contain a quadrilateral (that is a $K(2, 2)$). Denote by $d_1, d_2, \ldots, d_n$ the degrees of the vertices of G in the first vertex class, say V_1.

Then $\sum_1^n d_i = e = e(G) > z$ and, as in the proof of Lemma 7, we have

$$\binom{n}{2} \geq \sum_1^n \binom{d_i}{2} = \frac{1}{2}\sum_1^n d_i^2 - \frac{1}{2}\sum_1^n d_i$$
$$\geq \frac{1}{2n}e^2 - \frac{e}{2} > \frac{z(z-n)}{2n} = \binom{n}{2}.$$

This shows that the inequality does hold.

In fact, the proof so far tells us a considerable amount about the graphs G for which equality can be attained. We must have $d_1 = d_2 = \cdots = d_n = d$ and any two vertices in the second vertex class V_2 of G have exactly one common neighbour. By symmetry each vertex of V_2 has degree d and any two vertices in V_1 have exactly one common neighbour.

Call the vertices in V_1 *points* and the sets $\Gamma(x)$, $x \in V_2$, *lines*. By the remarks above there are n points and n lines, each point is on d lines and each line contains d points, there is exactly one line through any two points and any two lines meet in exactly one point. Thus we have arrived at the *projective plane* of order $d - 1$. Since the steps are easy to trace back, we see that equality holds for every n for which there is a projective plane with n points. In particular, equality holds for every $n = q^2 + q + 1$ where q is a prime power.

In conclusion, let us see the actual construction of G for the above values of n. Let q be a prime power and let $PG(2, q)$ be the projective plane over the field of order q. Let V_1 be the set of points and V_2 the set of lines. Then

$$|V_1| = |V_2| = q^2 + q + 1 = n.$$

Let G be the bipartite graph $G_2(n, n)$ with vertex classes V_1 and V_2 in which Pl ($P \in V_1$, $l \in V_2$) is an edge iff the point P is on the line l. (See Figure IV.4.) Then G has $n(q + 1) = \frac{1}{2}n\{1 + (4n - 3)^{1/2}\}$ edges and it does not contain a quadrilateral. □

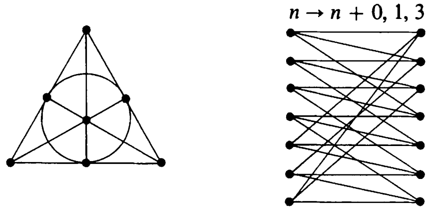


Figure IV.4. $PG(2, 2)$ and the corresponding bipartite graph.

§3 Hamilton Paths and Cycles

A class of graphs is said to be *monotone* if whenever a graph L belongs to the class and M is obtained from L by adding to it an edge (but no vertex) then M also belongs to the class. Most theorems in graph theory can be expressed by saying that a monotone class $\mathscr{M}$ is contained in a monotone class $\mathscr{P}$.

Of course, these classes are usually described in terms of graph invariants or subgraphs contained by them. For example, the simplest case of Turán's theorem, discussed in the previous section, states that the class $\mathcal{M} = \{G(n, m): m > n^2/4\}$ is contained in $\mathcal{P} = \{G: G \text{ contains a } K^3\}$. It is worth noting that a class $\mathcal{P}$ of graphs is said to be a *property of graphs* if $L \in \mathcal{P}$ and $L \cong M$ imply $M \in \mathcal{P}$.

How should we go about deciding whether or not $\mathcal{M}$ is contained in $\mathcal{P}$? In some cases there is a simple and beautiful way of tackling this problem. Suppose we have a class $\mathcal{T}$ of triples (G, x, y), where G is a graph and x and y are non-adjacent vertices of G, such that if $(G, x, y) \in \mathcal{T}$ and $G \in \mathcal{M}$ then G belongs to $\mathcal{P}$ iff $G^+ = G + xy$ does. This holds, for example, if $\mathcal{P}$ is the property of containing a K^r and $\mathcal{T} = \{(G, x, y): |\Gamma(x) \cap \Gamma(y)| < r - 2\}$. In this case G can be replaced by G^+. If G^+ also contains two non-adjacent vertices, say u and v, such that $(G^+, u, v) \in \mathcal{T}$ then we can repeat the operation: we can replace G^+ by $G^{++} = G^+ + uv$. Continuing in this way we arrive at a graph $G^* \supset G$ which belongs to $\mathcal{P}$ iff G does, and which is a *closure of G with respect to* $\mathcal{T}$, that is it has the additional property that for no vertices $a, b \in G^*$ does $(G^*, a, b) \in \mathcal{T}$ hold. Thus it is sufficient to decide about these graphs $G^* \in \mathcal{M}$ whether or not they belong to $\mathcal{P}$.

Of course the method above is feasible only if (i) the class $\mathcal{T}$ is simple enough, (ii) it is easy to show that G belongs to $\mathcal{P}$ iff G^+ does and (iii) if we start with a graph $G \in \mathcal{M}$ then a graph $G^* \in \mathcal{M}$ is easily shown to belong to $\mathcal{P}$. In this section we give two examples that satisfy all these requirements: we shall give sufficient conditions for a graph to contain a Hamilton cycle or a Hamilton path. Because of the special features of these examples it will be convenient to use slightly different notation and terminology.

Let n and k be natural numbers and let $\mathcal{P}$ be a class of graphs of order n. We say that $\mathcal{P}$ is *k-stable if* whenever G is an arbitrary graph of order n, x and y are non-adjacent vertices of G and $d(x) + d(y) \geq k$ then G has property $\mathcal{P}$ iff $G^+ = G + xy$ has it also. It is easily seen that for every graph G of order n there is *unique minimal graph* $G^* = C_k(G)$ *containing* G such that

$$d_{G^*}(x) + d_{G^*}(y) \leq k - 1 \quad \text{for} \quad xy \notin E(G^*).$$

In the notation of the previous paragraph, we shall take

$$\mathcal{T} = \{(G, x, y): |G| = n, xy \notin E(G), d(x) + d(y) \geq k\},$$

which is certainly simple enough, so (i) will be satisfied. It is also encouraging that $G^* = C_k(G)$ is unique. Almost by definition we have the following *principle of stability*: if $\mathcal{P}$ is a k-stable property of graphs of order n then G *has property $\mathcal{P}$ iff $C_k(G)$ has it also.* We call $C_k(G)$ the k-closure of G.

Requirement (ii) is also satisfied, since the gist of the proof of Theorem 2 is that the property of containing a Hamilton cycle is n-stable and the property of containing a Hamilton path is $(n - 1)$-stable. Indeed if $d(x) + d(y) \geq n - 1$ whenever x and y are non-adjacent distinct vertices, then the graph is connected so the proof of Theorem 2 can be applied. (In fact, this

is exactly what inspired the notion of a k-closure.) By the stability principle we obtain the following reformulation of Theorem 2 in the case $k = n$ or $n - 1$.

Lemma 11. *G is Hamiltonian iff $C_n(G)$ is and G has a Hamilton cycle iff so does $C_n(G)$.* □

Depending on the amount of work we are able and willing to put in at this stage (cf. requirement (iii)), we obtain various sufficient conditions for a graph to be Hamiltonian. Of course, the case $k = n$ of Theorem 2 is obtained without any work, and so is the case $k = n - 1$, since the conditions imply immediately that $C_n(G) = K^n$ in the first case and $C_{n-1}(G) = K^n$ in the second, and K^n is Hamiltonian if $n \geq 3$. In order to make better use of Lemma 11, we shall prove the following ungainly technical lemma.

Lemma 12. *Let k, n and t be natural numbers, $t < n$, and let G be a graph with vertex set $V(G) = \{x_1, x_2, \ldots, x_n\}$, whose k-closure $C_k(G)$ contains at most $t - 1$ vertices of degree $n - 1$. Then there are indices i, j, $1 \leq i < j \leq n$, such that $x_i x_j \notin E(G)$ and each of the following four inequalities holds:*

$$\begin{gathered} j \geq \max\{2n - k - i, n - t + 1\} \\ d(x_i) \leq i + k - n, \qquad d(x_j) \leq j + k - n - 1, \\ d(x_i) + d(x_j) \leq k - 1. \end{gathered} \tag{4}$$

Remark. Note that it is not assumed that the degree sequence $d(x_1)$, $d(x_2), \ldots, d(x_n)$ of G is ordered in any way.

Proof. The graph $H = C_k(G)$ is not complete so we can define two indices i and j as follows:

$$\begin{aligned} j &= \max\{l\colon d_H(x_L) \neq n - 1\}, \\ i &= \max\{l\colon x_l x_j \notin E(H)\}. \end{aligned}$$

Then $x_i x_j \notin E(H)$ so, by (3),

$$d_H(x_i) + d_H(x_j) \leq k - 1,$$

which implies the fourth inequality in (4). Each of the vertices

$$x_{j+1}, x_{j+2}, \ldots, x_n$$

has degree $n - 1$ in H so

$$n - j \leq t - 1$$

and

$$n - j \leq \delta(H) \leq d_H(x_i).$$

The vertex x_j is joined to the $n - j$ vertices following it and to the $j - i - 1$ vertices preceding it, so

$$d_H(x_j) \geq n - j + j - i - 1 = n - i - 1.$$

These inequalities enable us to show that the indices i, j, $1 \leq i < j \leq n$, satisfy the remaining three inequalities in (4). Indeed,

$$d_G(x_i) \leq d_H(x_i) \leq k - 1 - d_H(x_j) \leq k - 1 - (n - i - 1) = i + k - n,$$

$$d_G(x_i) \leq d_H(x_j) \leq k - 1 - d_H(x_i) \leq k - 1 - (n - j) = j + k - n - 1,$$

$$i + j \geq (n - d_H(x_j) - 1) + (n - d_H(x_i)) \geq 2n - 1 - (k - 1) = 2n - k.$$

□

Combining Lemma 11 and Lemma 12 (with $t = n - 1$ and $k = n$ or $n - 1$) we obtain rather complicated but useful conditions for the existence of a Hamilton path or cycle.

Theorem 13. *Let G be a graph with vertex set $V(G) = \{x_1, x_2, \ldots, x_n\}$, $n \geq 3$. Let $\varepsilon = 0$ or 1 and suppose there are no indices i, 1, $1 \leq i < j \leq n$, such that $x_i x_j \notin E(G)$ and*

$$j \geq n - i + \varepsilon,$$

$$d(x_i) \leq i - \varepsilon, \qquad d(x_j) \leq j - 1 - \varepsilon,$$

$$d(x_i) + d(x_j) \leq n - 1 - \varepsilon.$$

If $\varepsilon = 0$ then G has a Hamilton cycle and if $\varepsilon = 1$ then G has a Hamilton path.

□

This theorem has the following beautiful consequence.

Corollary 14. *Let G be a graph with degree sequence $d_1 \leq d_2 \leq \cdots \leq d_n$, $n \geq 3$, and let $\varepsilon = 0$ or 1. Suppose*

$$d_{n-k-\varepsilon} \geq n - k \quad \textit{whenever } d_k \leq k - \varepsilon < \tfrac{1}{2}(n - \varepsilon).$$

If $\varepsilon = 0$ then G has a Hamilton cycle and if $\varepsilon = 1$ then G has a Hamilton path.

□

We draw the attention of the reader to Exercises 21 and 22 which show that the assertions in the corollary above are in some sense best possible. In particular, if $d_1 \leq d_2 \leq \cdots \leq d_n$ is a graphic sequence such that

$$d_{n-k} < n - k \quad \text{and} \quad d_k \leq k < \frac{n}{2}$$

then there is a graph G with vertex set $\{x_1, x_2, \ldots, x_n\}$ such that $d(x_i) \geq d_i$, $1 \leq i \leq n$, and G does not have a Hamilton cycle.

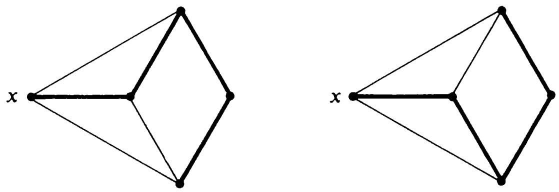


Figure IV.5. An x-path and a simple transform of it.

There is another customary way of showing that a graph has a Hamilton cycle or path. Let S be a longest x_0-*path* in G, that is a longest path beginning at $x_0: S = x_0 x_1 \dots x_k$. Then $\Gamma(x_k) \subset \{x_0, x_1, \dots, x_{k-1}\}$ since otherwise S could be continues to a longer path. If x_k is adjacent to x_j, $1 \le j < k - 1$, then $S' = x_0 x_1 \dots x_j x_k x_{k-1} \dots x_{j+1}$ is another longest x_0-path. We call S' a *simple transform* of S. It is obtained from S by erasing the edge $x_j x_{j+1}$ and adding to it the edge $x_k x_j$. Note that if S' is a simple transform of S then S is a simple transform of S' and S has exactly $d(x_k) - 1$ simple transforms. The result of a sequence of simple transforms is called a transform (See Figure IV.5).

Let L be the set of endvertices (different from x_0) of transforms of S and put $N = \{x_j \in S: x_{j-1} \in L \text{ or } x_{j+1} \in L\}$ and $R = V - N \cup L$. Thus L is the collection of the last vertices of the transforms, N is the collection of their neighbours on S and R is the rest of the vertices.

Theorem 15. *The graph G has no L-R edges.*

Proof. Recall that there is no $L - (V(G) - V(S))$ edge, since S is a longest x_0-path, so in particular $V(S) = V(P)$ for every transform P of S.

Suppose $x_i x_j \in E(G)$, where $x_i \in L$ and $x_j \in R$. Let S_i be a transform of S ending in x_i. Since at least one neighbour of x_j on S_i is the endvertex of a simple transform of S_i, x_j cannot have the same neighbours on S and S_i, since otherwise x_j would belong to N. However, when the edge $x_{j'} x_{j'}$ $j' = j - 1$ or $j + 1$, is erased during a sequence $S \to S' \to S'' \to \cdots \to S_i$ of simple transformations, one of the vertices $x_{j'}$, x_j is put into L and the other into N. Thus $x_j \in L \cup N = V(G) - R$, contradicting our assumption. □

We will make use of Theorem 15 later in Chapter VII. The last two results of this section are also obtained with the use of simple transforms.

Theorem 16. *Let W be the set of vertices of even degree in a graph G and let x_0 be a vertex of G. Then there are an even number of longest x_0-paths ending in W.*

Proof. Let H be the graph whose vertex set is the set $\sum$ of longest x_0-paths in G, in which $P_1 \in \sum$ is joined to $P_2 \in \sum$ iff P_2 is a simple transform of P_1. Since the degree of $P = x_0 x_1 \dots x_k \in \sum$ in H is $d(x_k) - 1$, the set of longest paths ending in W is exactly the set of vertices of odd degree in H. The

number of vertices of odd degree is even in any graph, so the proof is complete. □

Theorem 17. *Let G be a graph in which every vertex has odd degree. Then every edge of G is contained in an even number of Hamilton cycles.*

PROOF. Let $x_0 y \in E(G)$. Then in $G' = G - x_0 y$ only x_0 and y have even degree, so in G' there are an even number of longest x_0-paths that end in y. Thus either G has no Hamilton cycle that contains $x_0 y$ or it has a positive even number of them. □

§4 The Structure of Graphs

As we saw in §2 on the example of a Turán graph $T_{r-1}(n)$, a graph of order n with $\frac{1}{2}((r-2)/(r-1)+o(1))n$ edges does not necessarily contain a K^r, a complete graph of order r. (Here and in what follows we use Landau's notation $o(1)$ to denote a function tending to 0 as $n \to \infty$.) The main aim of this section is to prove the deep result of Erdös and Stone, proved in 1946, that if $\varepsilon > 0$ then εn^2 more edges ensure not only a K^r, but a $K_r(t)$, a complete r-partite graph with t vertices in each class, where $t \to \infty$ as $n \to \infty$.

A sharper version of the theorem is due to Bollobás and Erdös; this is the result we shall present here. The simple proof is based on two lemmas. The first lemma enables us to replace a condition on the size by a condition on the minimum degree; the second is a technical lemma in the vein of Lemma 7.

Lemma 18. *Let $0 < a < a + \varepsilon < 1$ and put $\eta = (\varepsilon/2)^{1/2}$. Then every graph G of order $n \geq 2a/\varepsilon$ and size at least $\frac{1}{2}(a+\varepsilon)n^2$ contains a subgraph H with $|H| = h \geq \eta n$ and minimum degree at least ah.*

PROOF. Define a sequence of graphs $G_0 = G \supset G_1 \supset G_2 \supset \ldots$ with $|G_k| = n - k$ as follows. If G_k has a vertex x_k of degree less than $a(n-k)$ then put $G_{k+1} = G_k - x_k$, if $\delta(G_k) \geq a(n-k)$ then terminate the sequence. The assertion of the theorem follows if we show that the sequence has to terminate in a graph G_k with $h = n - k \geq \eta n$.

Suppose this is not so and we arrive at a graph G_k with $h = n - k < \eta n$. By construction

$$e(G_k) \geq \tfrac{1}{2}(a+\varepsilon)n^2 - a\{n + (n-1) + \cdots + (h+1)\}.$$

On the other hand

$$e(G_k) \leq e(K^h) = \binom{h}{2}$$

so

$$\varepsilon n^2 - an + ah + h \le (1 - a)h^2,$$

contradicting our assumptions. □

Lemma 19. *Let $r \ge 2, 1 \le t \le q$ and $N = n - (r - 1)q \ge 1$. Let G be a graph of order n that contains a $K_{r-1}(q)$, say $\tilde{K}$, but does not contain a $K_r(t)$. Then G has at most*

$$e = ((r - 2)q + t)N + 2qN^{1-(1/t)}$$

edges of the form xy, where $x \in \tilde{K}$ and $y \in G - \tilde{K}$.

Proof. Let $S_1, S_2, \ldots, S_{r-1}$ be the vertex classes of $\tilde{K}$ and put $S = \bigcup_1^{r-1} S_i$. Thus $|S_i| = q$ and $|S| = (r - 1)q$. Let $x_1, x_2, \ldots, x_N$ be the vertices of $G - \tilde{K}$ and put $A_i = \Gamma(x_i) \cap S, i = 1, 2, \ldots, N$. With this notation the lemma states that

$$\sum_1^N |A_i| \le e.$$

Call a set $W \subset S$ a *centipede* if it has exactly t vertices in each set S_i. Since G does not contain a $K_r(t)$, every centipede is contained in at most $t - 1$ of the A_i. We may assume that $|A_i| \ge (r - 2)q + t$ for $i \le M$ and $|A_i| < (r - 2)q + t$ for $i > M$. For $i \le M$ and $j \le r - 1$ put $a_{ij} = |A_i \cap S_j|$. Then $a_{ij} \ge t$ and A_i contains

$$\prod_{j=1}^{r-1} \binom{a_{ij}}{t}$$

centipedes. Since altogether there are $\binom{q}{t}^{r-1}$ centipedes and each one of them is contained in at most $t - 1$ sets A_i, we have

$$\sum_{i=1}^{M} \prod_{j=1}^{r-1} \binom{a_{ij}}{t} \le (t - 1)\binom{q}{t}^{r-1}.$$

The required estimate follows from this inequality and the convexity of the binomial coefficients. Indeed putting $a_i = \sum_{j=1}^{r-1} a_{ij} - (r - 2)q$, one has

$$\prod_{j=1}^{r-1} \binom{a_{ij}}{t} \ge \binom{q}{t}^{r-2} \binom{a_i}{t},$$

so

$$\sum_{i=1}^{M} \binom{a_i}{t} \le (t - 1)\binom{q}{t}.$$

As $\binom{x}{t}$ is a convex function of x, on putting $a = \sum_1^M a_i/M$ the last inequality implies

$$M\binom{a}{t} \le (t - 1)\binom{q}{t}.$$

Thus

$$a \leq (t - 1)^{1/t} q M^{-1/t} + t < 2qM^{-1/t} + t$$

and

$$\sum_{1}^{M} |A_i| < 2qM^{-1/t} + M((r - 2)q + t).$$

Since

$$\sum_{M+1}^{N} |A_i| \leq (N - M)((r - 2)q + t),$$

the required estimate does hold. □

Theorem 20. *Let $r \geq 2$ be an integer and let $0 < \varepsilon < \frac{1}{2}(r - 1)$. Then there exists a $d = d(\varepsilon, r) > 0$ such that if n is sufficiently large and*

$$m \geq \{\tfrac{1}{2}(r - 2)/(r - 1) + \varepsilon\} n^2$$

then every graph of order n and size m contains a $K_r(t)$ with $t \geq \lfloor d \log n \rfloor$.

Proof. We shall apply induction on r. For $r = 2$ Theorem 9 implies that any $d < -1/\log(2\varepsilon)$ will do for $d(\varepsilon, 2)$. Indeed, if there is a graph $G(n, m)$ without a $K_2(t)$ where $m \geq \varepsilon n^2$ and $t = \lfloor d \log n \rfloor$, then by Theorem 9 we have

$$\varepsilon n^2 \leq \tfrac{1}{2}(t - 1)^{1/t} n^{2-(1/t)} + \tfrac{1}{2} t n,$$

that is

$$2\varepsilon n^{1/t} \leq (t - 1)^{1/t} + t n^{-1+(1/t)}.$$

However, this inequality cannot hold if n is sufficiently large.

Now assume that $r > 2$ and the result holds for smaller values of r. For the sake of convenience we write $K_p(x)$ for $K_p(\lfloor x \rfloor)$ and we denote by $G_\varepsilon(n)$ a graph of order n and size *at least* $\{\frac{1}{2}(r - 2)/(r - 1) + \varepsilon\} n^2$.

Put $\varepsilon_r = \frac{1}{2}\{(r - 2)/(r - 1) - (r - 3)/(r - 2)\} = \{2(r - 1)(r - 2)\}^{-1} > 0$ and $e_r = d(\varepsilon_r, r - 1) > 0$. Then e_r depends only on r, so by the induction hypothesis if n is large then every $G_\varepsilon(n)$ contains a $K_{r-1}(e_r \log n) = K_{r-1}(q)$. We shall show that $d = d(\varepsilon, r) = e_r(r - 1)\varepsilon/4$ will do for the theorem, that is *every $G_\varepsilon(n)$* contains a $K_r(d \log n) = K_r(t)$ *if n is large enough.*

Suppose that, contrary to this assertion, there are arbitrarily large values of n for which some $G_\varepsilon(n)$ does not contain a $K_r(d \log n) = K_r(t)$. Given any $h_0 \geq 1/\varepsilon$, let $G = G_\varepsilon(n)$, $n \geq h_0/\eta$, contain a $K_{r-1}(q)$, say $\tilde{K}$, but no $K_r(t)$. Hence by Theorem 9 there are at most $\frac{1}{2}(t - 1)q^{2-(1/t)} + \frac{1}{2}qt$ edges in a subgraph spanned by the q vertices in a class of $\tilde{K}$, so there are at most

$$A = \binom{r-1}{2} q^2 + (r - 1)\{\tfrac{1}{2}(t - 1)q^{2-1/t} + \tfrac{1}{2}qt\}$$

edges in the subgraph spanned by the whole of $\tilde{K}$. Furthermore, by Lemma 19, at most $B = ((r-2)q + t)(n-(r-1)q) + 2qn^{1-(1/t)}$ edges join $\tilde{K}$ to $G - \tilde{K}$. Finally, as $\delta(G) \geq \{(r-2)/(r-1) + \varepsilon\}n$, we must have

$$2A + B \geq (r-1)q\delta(G) \geq \{(r-2) + \varepsilon(r-1)\}qn,$$

and so

$$2qn^{1-(1/t)} + (r-1)q^2 + tn \leq \varepsilon(r-1)qn.$$

We have arrived at a contradiction, since this inequality does not hold for larger values of n. □

Remark. In a certain sense Theorem 20 is best possible: for every ε and r there is a constant d_r^* tending to 0 with ε such that the graph described in the theorem need not contain a $K_r(t)$ with $t = \lfloor d_r^* \log n \rfloor$. In fact, we shall see in Chapter VII Theorem 3 that if $0 < \varepsilon < \frac{1}{2}$ and $d_2^* > -2/\log(2\varepsilon)$, then for every sufficiently large n there is a graph $G(n, m)$ not containing a $K_2(t)$, where $m = \lfloor \varepsilon n^2 \rfloor$ and $t = \lfloor d_2^* \log n \rfloor$. This result will imply immediately (cf. Exercise 6 of Chapter VII) that if $r > 2$ and $0 < \varepsilon < \frac{1}{2}(r-1)^{-2}$ then any value greater than $-2/\log(2(r-1)^2\varepsilon)$ will do for d_r^*.

Since $d \log n \to \infty$ as $n \to \infty$, Theorem 20 has the following immediate corollaries.

Corollary 21. *Let $F = K_r(t)$, where $r \geq 2$ and $t \geq 1$. Then the maximal size of a graph of order n without a $K_r(t)$ is*

$$\mathrm{ex}(n; F) = \frac{1}{2}\left\{\frac{r-2}{r-1} + o(1)\right\}n^2. \qquad \square$$

Corollary 22. *Let $F_1, F_2, \ldots, F_l$ be non-empty graphs. Denote by r the minimum of the chromatic numbers of the F_i, that is let r be the minimum number for which at least one of the F_i is contained in an $F = K_r(t)$ for some t. Then the maximal size of a graph of order n not containing any of the F_i is*

$$\mathrm{ex}(n; F_1, F_2, \ldots, F_l) = \frac{1}{2}\left\{\frac{r-2}{r-1} + o(1)\right\}n^2.$$

PROOF. The Turán graph $T_{r-1}(n)$ does not contain any of the F_i, so

$$\mathrm{ex}(n; F_1, F_2, \ldots, F_l) \geq e(T_{r-1}(n)) = t_{r-1}(n) = \frac{1}{2}\left\{\frac{r-2}{r-1} + o(1)\right\}n^2.$$

Conversely, since, say $F_j \subset F = K_r(t)$ for some j and t,

$$\mathrm{ex}(n; F_1, F_2, \ldots, F_l) \leq \mathrm{ex}(n; F_j) \leq \mathrm{ex}(n; F) = \frac{1}{2}\left\{\frac{r-2}{r-1} + o(1)\right\}n^2. \qquad \square$$

Theorem 20 is the basis of a rather detailed study of the structure of extremal graphs, giving us considerably more accurate results than Corollary 22. This theory is, however, outside the scope of our book.

The *density* of a graph G of order n is defined to be $e(G)/\binom{n}{2}$. The *upper density* of an *infinite* graph G is the supremum of the densities of arbitrarily large *finite* subgraphs of G. It is surprising and fascinating that not every value between 0 and 1 is the upper density of some infinite graph; in fact, the range of the upper density is a countable set.

Corollary 23. *The upper density of an infinite graph G is $1, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots,$ or 0. Each of these values is the upper density of some infinite graph.*

PROOF. Let G_r be the complete r-partite graph with infinitely many vertices in each class. Since the density of $K_r(t)$ tends to $1 - (1/r)$ as t tends to ∞, the upper density of G_r is $1 - (1/r)$, proving the second assertion.

Now let α be the upper density of G and suppose

$$\alpha > 1 - \frac{1}{r-1} \qquad (r \geq 2).$$

Then there is an $\varepsilon > 0$ such that G contain graphs H_k of order n_k with $n_k \to \infty$ satisfying $e(H_k) \geq \frac{1}{2}(1 - (1/(r-1)) + \varepsilon)n_k^2$. By Theorem 19 each H_k contains a subgraph $K_r(t_k)$ with $t_k \to \infty$; the subgraphs $K_r(t_k)$ show that $\alpha \geq 1 - (1/r)$. □

EXCERCISES

1. (i) Let G be a graph of order n, maximum degree $\Delta \geq 3$ and diameter d. Let $n_0(g, \delta)$ be as in Theorem 1. Prove that $n \leq n_0(2d + 1, \Delta)$ with equality iff G is Δ-regular and has girth $2d + 1$.
   (ii) Let G be a graph of order n, maximum degree $\Delta \geq 3$, and suppose every vertex is within distance $d - 1$ of each pair of adjacent vertices. Prove that $n \leq n_0(2d, \Delta)$, with equality iff G is Δ-regular and has girth $2d$.
2. Prove Theorem 4 for $k = 3$ and 4.
3. Prove that the maximal number of edges in a graph of order n without an even cycle is $\lfloor \frac{3}{2}(n-1) \rfloor$. Compare this with the maximum size of a graph without an odd cycle.
4. Show that a tree with $2k$ endvertices contains k edge-disjoint paths joining distinct endvertices.
5. Suppose x is not a cutvertex and has degree $2k$. Prove that there are k edge-disjoint cycles containing x. [Cf. Exercise 4.]
6.⁻ Show that a graph with n vertices and minimum degree $\lfloor (r-2)n/(r-1) \rfloor + 1$ contains a K^r.
7. Let G have $n \geq r + 1$ vertices and $t_{r-1}(n) + 1$ edges.
   (i) Show that G contains two K^r subgraphs with $r - 1$ vertices in common.
   (ii) Prove that for every p, $r \leq p \leq n$, G has a subgraph with p vertices and at least $t_{r-1}(p) + 1$ edges.

8. Prove that for $n \geq 5$ every graph of order n with $\lfloor n^2/4 \rfloor + 2$ edges contains two triangles with exactly one vertex in common.

9. Prove that if a graph with n vertices and $\lfloor n^2/4 \rfloor - l$ edges contains a triangle then it contains at least $\lfloor n/2 \rfloor - l - 1$ triangles. [*Hint.* Let $x_1x_2x_3$ be a triangle and denote by m the number of edges joining $\{x_1, x_2, x_3\}$ to $V(G) - \{x_1, x_2, x_3\}$. Estimate the number of triangles in $G - \{x_1, x_2, x_3\}$ and the number of triangles sharing a side with $x_1x_2x_3$.]

10.$^+$ (i) Show that the edges of a graph of order n can be covered with not more than $\lfloor n^2/4 \rfloor$ edges and triangles.
(ii) Let G be a graph with vertices $x_1, x_2, \ldots, x_n$, $n \geq 4$. Prove that there is a set S, $|S| \leq \lfloor n^2/4 \rfloor$, containing non-empty subsets $X_1, X_2, \ldots, X_n$ such that x_ix_j is an edge of G if $X_i \cap X_j \neq \emptyset$.

11.$^-$ Let $1 \leq k < n$. Show that every graph of order n and size $(k-1)n - \binom{k}{2} + 1$ contains a subgraph with minimum degree k, but there is a graph of order n and size $(k-1)n - \binom{k}{2}$ in which every subgraph has minimum degree at most $k - 1$. [*Hint.* Imitate the proof of Lemma 18.]

12.$^-$ Show that a graph of order n and size $(k-1)n - \binom{k}{2} + 1$ contains every tree of order $k + 1$.

13. Let G be a graph of order n which does not contain a cycle and one of its diagonals. Prove that if $n \geq 4$ then G has at most $2n - 4$ edges.
Show that if $n \geq 6$ and G has $2n - 4$ edges then G is the complete bipartite graph $K(2, n-2)$. [*Hint.* Consider a longest path in G.]

14.$^-$ Let $k \geq 1$ and let G be a graph of order n without an odd cycle of length less than $2k + 1 \geq 5$. Prove that $\delta(G) \leq \lfloor n/2 \rfloor$ and $T_2(n)$ is the only extremal graph, unless $n = 2k + 1 = 5$, in which case there is another extremal graph, C^5.

15.$^+$ Let G be a graph of order n without an odd cycle of length less than $2k + 1 \geq 5$. Prove that if G does not contain $\lceil n/2 \rceil$ independent vertices then $\delta(G) \leq 2n/(2k+1)$. Show that equality holds only for $n = (2k+1)t$ and the extremal graphs are obtained from a cycle C^{2k+1} by replacing each vertex by t vertices, as in Figure IV.6.

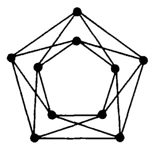


Figure IV.6. The graph $C^5(2)$.

16. Let $x_1, x_2, \ldots, x_n$ be vectors of norm at least 1 in a Euclidean space. Prove that there are at most $\lfloor n^2/4 \rfloor$ unordered pairs i, j such that $|x_i + x_j| < 1$. [*Hint.* Show that if $|x_1| = |x_2| = |x_3| = 1$ then $|x_i + x_j| \geq 1$ for some i, j, $1 \leq i < j \leq 3$.]

17.$^+$ Let X and Y be independent identically distributed random variables taking values in a Euclidean space. Prove that $P(|X + Y| \geq x) \geq \frac{1}{2}P(|X| \geq x)^2$ for every $x \geq 0$.

18. Let $x_1, x_2, \ldots, x_{3p} \in \mathbb{R}^2$ be such that $|x_i - x_j| \le 1$. Prove that at most $3p^2$ of the distances $|x_i - x_j|$ are greater than $\sqrt{2}/2$. [*Hint.* Show that among any four of the points there are two within distance $\sqrt{2}/2$ of each other.]

19. The *clique number* $\mathrm{cl}(G)$ of a graph G is the maximum order of a complete subgraph of G. Show that if a regular graph of order n contains a complete subgraph of order $\lfloor n/2 \rfloor + 1$ then it is a complete graph, but for every p, $1 \le p \le n/2$, there is a regular graph of order n with $\mathrm{cl}(G) = p$.

20.⁻ We say that a set $W \subset V(G)$ *covers* the edges of a graph G if every edge of G is incident with at least one vertex in W. Denote by $\alpha_0(G)$ the minimum number of vertices covering the edges of G. Prove that if G has n vertices and m edges then $\alpha_0(G) \le 2mn/(2m + n)$, with equality iff $G = pK^r$ for some p and r, that is iff each component of G is a K^r for some r. [*Hint.* Note that $\alpha_0(G) = n - \mathrm{cl}(\bar{G})$, and if $\mathrm{cl}(\bar{G}) = p$ then by Turán's theorem $e(\bar{G}) \le t_p(n)$, so $m = \binom{n}{2} - e(\bar{G}) \ge \binom{n}{2} - t_p(n)$.]

21. (Cf. Corollary 14.) Let $d_1 \le d_2 \le \cdots \le d_n$ be a graphic sequence such that for some k

$$d_k \le k < \frac{n}{2} \quad \text{and} \quad d_{n-k} \le n - k - 1.$$

Show that there is a non-Hamiltonian graph G with vertex set $\{x_1, x_2, \ldots, x_n\}$ such that $d(x_i) \ge d_i$, $1 \le i \le n$ (cf. Figure IV.7).



Figure IV.7. The graph $(K^3 \cup E^3) + K^3$ has no Hamilton cycle and $(K^2 \cup E^3) + K^2$ has no Hamilton path.

22. (Cf. Corollary 14) Let $d_1 \le d_2 \le \cdots \le d_n$ be a graphic sequence such that for some k

$$d_k \le k - 1 < \tfrac{1}{2}(n - 1) \quad \text{and} \quad d_{n+1-k} \ge n - k.$$

Prove that there is a graph G with vertex set $\{x_1, x_2, \ldots, x_n\}$ such that $d(x_i) \ge d_i$, $1 \le i \le n$, and G does not contain a Hamilton path (cf. Figure IV.7).

23. Prove that a non-Hamiltonian graph of order $n \ge 3$ has at most $\binom{n}{2} - (n - 2)$ edges and there is a unique extremal graph.

    Prove that a graph of order $n \ge 2$ without a Hamilton path has at most $\binom{n}{2} - (n - 3)$ edges and $K^{n-1} \cup E^1$ is the unique extremal graph.

24. Given $\delta < n/2$, determine the maximal number of edges in a graph G of order n without a Hamilton cycle (path), provided $\delta(G) = \delta$.

25. Prove Theorem 4 by making use of simple transforms of a longest x_0-path $P = x_0 x_1 \ldots x_t$. [*Hint.* Apply induction on n. If $\delta(G) \le k/2$, the result follows by induction, otherwise consider the set L of endvertices of simple transforms of P. Put $l = |L|$, $r = \max_{x \in L} d(x)$ and note that $l \ge r$ and the neighbours of each $x \in L$ are contained in $\{x_t, x_{t-1}, \ldots, x_{t-k+1}\}$. Deduce that $e(G) - e(G - L) \le l(k - l) + l(r + l - k) \le kl/2$ and complete the proof by applying the induction hypothesis to $G - L$.]

26. Let $1 < a_1 < a_2 < \cdots < a_k \le x$ be natural numbers. Suppose no a_i divides the product of any two others. Prove that $k \le \pi(x) + x^{2/3}$, where, as usual, $\pi(x)$ denotes the number of primes not exceeding x. [*Hint.* Put $V_1 = \{1, 2, \ldots, \lfloor x^{2/3} \rfloor\}$ and $V_2 = \{x: x^{2/3} \le b \le x$ and b is a prime$\}$. Show first that $a_i = b_i c_i$, where $b_i, c_i \in V = V_1 \cup V_2$. Let G be the graph (with loops) with vertex set V whose edges (loops) are $b_i c_i$. Note that G does not contain a path of length 3.]

27.$^+$ (Cf. Exercise 26). Let $1 < a_1 < a_2 < \cdots < a_k \le x$ be natural numbers. Suppose $a_i a_j \ne a_k a_l$ unless $\{i, j\} = \{h, l\}$. Prove that $k \le \pi(x) + cx^{3/4}$ for some constant $c > 0$. [*Hint.* The graph G of Exercise 26 contains no quadrilaterals; apply Theorem 8 to the bipartite subgraph of G with vertex classes V_1 and V_2. Recall the prime number theorem, namely that $(\pi(x)\log x)/x \to 1$ as $x \to \infty$.]

28. Denote by $D_k(n)$ the maximum number of occurrences of the same positive distance among n points in $\mathbb{R}^k$. Prove that $\lim_{n\to\infty} D_k(n)/n^2 = \frac{1}{2} - (1/2\lfloor k/2 \rfloor)$ if $k \ge 2$. [*Hint.* (i) Note that if $\mathbf{x} \in \{\mathbf{z} \in \mathbb{R}^k: z_1^2 + z_2^2 = 1$ and $z_i = 0$ if $i > 2\}$ and $\mathbf{y} \in \{\mathbf{z} \in \mathbb{R}^k: z_3^2 + z_4^2 = 1$ and $z_i = 0$ if $i \ne 3$ or $4\}$ then $|\mathbf{x} - \mathbf{y}| = \sqrt{2}$. (ii) Deduce from Theorem 20 that $D_k(n)$ is at least as large as claimed.]

Notes

The results concerning Hamilton cycles and paths presented in the chapter all originate in a paper of G. A. Dirac: Some theorems on abstract graphs, *Proc. London Math. Soc.* **2** (1952) 69–81. The notion of the k-closure discussed in §3 was introduced in J. A. Bondy and V. Chvátal, A method in graph theory, *Discrete Math.* **15** (1976) 111–135; Corollary 14, characterizing minimal forcibly Hamiltonian degree sequences, is in V. Chvátal, On Hamilton's ideals, *J. Combinatorial Theory* Ser. B **12** (1972) 163–168. The special case of Theorem 17 concerning cubic (that is 3-regular) graphs was proved by C. A. B. Smith, and was first mentioned in W. T. Tutte, On Hamiltonian circuits, *J. London Math. Soc.* **21** (1946) 98–101; the theorem itself is due to A. G. Thomason, Hamiltonian cycles and uniquely edge colourable graphs, *Annals of Discrete Math.* **3** (1978) 259–268.

The fundamental theorem of Turán concerning complete subgraphs is in P. Turán, On an extremal problem in graph theory (in Hungarian), *Mat. és Fiz. Lapok* **48** (1941) 436–452; its extension, Theorem 6, giving more information about the degree sequence of a graph without a K^r is in P. Erdös, On the graph theorem of Turán (in Hungarian), *Mat. Lapok* **21** (1970) 249–251. The Erdös–Stone theorem is in P. Erdös and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946) 1087–1091, its extension, Theorem 20, is in B. Bollobás and P. Erdös, On the structure of edge graphs, *Bull. London Math. Soc.* **5** (1973) 317–321. A rather detailed exposition of the results concerning the structure of extremal graphs, referred to at the end of §4, can be found in Chapter VI of B. Bollobás, *Extremal Graph Theory*, Academic Press, London—New York—San Francisco, 1978; Chapter III of the same book concerns cycles and contains many more results than we could present in this chapter.

CHAPTER V
Colouring

We wish to arrange the talks in a congress in such a way that no participant will be forced to miss a talk he would like to hear, that is that there are no undesirable clashes. Assuming a good supply of lecture rooms enabling us to hold as many parallel talks as we like, how long will the program have to last? Let us reformulate this question in terms of graphs. Let G be the graph whose vertices are the talks and in which two talks are joined iff there is a participant wishing to attend both. What is the minimum value of k for which $V(G)$ can be partitioned into k classes, say $V_1, V_2, \ldots, V_k$, such that no edge joins two vertices of the same class? As in §4 Chapter IV, we denote this minimum by $\chi(G)$ and call it the (*vertex*) *chromatic* number of G. The terminology originates in the usual definition of $\chi(G)$: a *colouring* of the vertices of G is an assignment of colours to the vertices in such a way that adjacent vertices have distinct colours; $\chi(G)$ is then the minimal number of colours in a (vertex) colouring of G.

Let us remark here that we shall use real colours (red, blue, ...) only if there are few colours, otherwise the natural numbers will be our "colours". Thus a k-colouring of the vertices of G is a function $c: V(G) \to \{1, 2, \ldots, k\}$ such that each set $c^{-1}(j)$ is independent.

Another scheduling problem goes as follows. Each of n businessmen wishes to hold confidential meetings with some of the others. Assuming that each meeting lasts a day and at each meeting exactly two businessmen are present, in how many days can the meetings be over? In this case one considers the graph H whose vertices correspond to the n businessmen and two vertices are adjacent iff the two businessmen wish to hold a meeting. Then the problem above asks for the minimal number of colours in an edge-colouring of H, that is in a colouring of the edges of H in such a way that no

two adjacent edges have the same colour. This number, denoted by $\chi'(H)$, is the *edge chromatic number* of H.

In this chapter we shall present some of the basic properties of the vertex and edge chromatic numbers. The final section (§3) is devoted to colouring planar graphs and to a very brief outline of the proof of the best known result in graph theory, the so called four colour theorem.

§1 Vertex Colouring

In Chapter I. §2 we noted the trivial fact that a graph is bipartite iff it does not contain an odd cycle. Thus $\chi(G) \geq 2$ iff G contains an edge and $\chi(G) \geq 3$ iff G contains an odd cycle. For $k \geq 4$ we do not have a similar characterization of graphs with chromatic number at least k, though there are some complicated characterizations (cf. Exercises 32–35). Of course, if $G \supset K^k$ then $\chi(G) \geq k$ and if G does not contain $h + 1$ independent vertices then $\chi(G) \geq |G|/h$. However, it is not easy to see that there are triangle-free graphs of large chromatic number (cf. Exercise 11), though in Chapter VII we shall show that there exist graphs with arbitrarily large chromatic number and arbitrarily large girth. The difficulty we encounter in finding such graphs shows that it would be unreasonable to expect a simple characterization of graphs with large chromatic number. Thus we shall concentrate on finding efficient ways of colouring a graph.

How should one try to colour the vertices of a graph with colours $1, 2, \ldots,$ using as few colours as possible? A simple way is as follows. Order the vertices, say $x_1, x_2, \ldots, x_n$ and then colour them one by one: colour x_1 1, then colour x_2 1 if $x_1x_2 \notin E(G)$ and 2 otherwise, and so on, colour each vertex the smallest colour it can have at that stage. This so-called *greedy algorithm* does produce a colouring, but this colouring may and usually does use many more colours than necessary. Figure V.I shows a bipartite (i.e., 2-colourable) graph for which the greedy algorithm wastes four colours. However, it is easily seen (Exercise 3′) that for every graph the vertices can be ordered in such a way that the greedy algorithm uses as few colours as possible. Thus it is not surprising that it pays to investigate the number of colours needed by the greedy algorithm in various orders of the vertices.

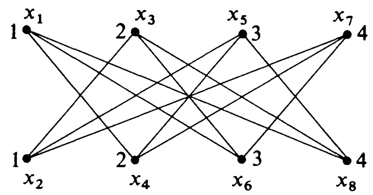


Figure V.1. In the order $x_1, x_2, \ldots, x_8$ the greedy algorithm needs four colours.

Let us start with an immediate consequence of the algorithm.

Theorem 1. *Let* $k = \max_H \delta(H)$, *where the maximum is taken over all spanned subgraphs of G. Then* $\chi(G) \le k + 1$.

PROOF. Write x_n for a vertex of degree at most k, and put $H_{n-1} = G - \{x_n\}$. By assumption H_{n-1} has a vertex of degree at most k. Let x_{n-1} be one of them and put $H_{n-2} = H_{n-1} - \{x_{n-1}\} = G - \{x_n, x_{n-1}\}$. Continuing in this way we enumerate all the vertices.

Now the sequence $x_1, x_2, \ldots, x_n$ is such that each x_j is joined to at most k vertices preceding it. Hence the algorithm will never need colour $k + 2$ to colour a vertex. □

It is, of course, very easy to improve the efficiency of the greedy algorithm. If we already know a subgraph H_0 which can be coloured with few colours, in particular, if we know a colouring of H_0 with $\chi(H_0)$ colours, then we may start our sequence with the vertices of H_0, colour H_0 in an efficient way and apply only then the algorithm to colour the remaining vertices. This gives us the following extension of Theorem 1.

Theorem 2. *Let* H_0 *be a spanned subgraph of G and suppose every subgraph H satisfying* $H_0 \subset H \subset G$, $V(H_0) \ne V(H)$, *contains a vertex* $x \in V(H) - V(H_0)$ *with* $d_H(x) \le k$. *Then*

$$\chi(G) \le \max\{k + 1, \chi(H)\}. \qquad \square$$

In some cases the problem of colouring a graph can be reduced to the problem of colouring certain subgraphs of it. This happens if the graph is disconnected or has a cutvertex or, slightly more generally, contains a complete subgraph whose vertex set disconnects the graph. Then we may colour each part separately since, at worst by a change of notation, we can fit these colourings together to produce a colouring of the original graph, as shown in Figure V.2.

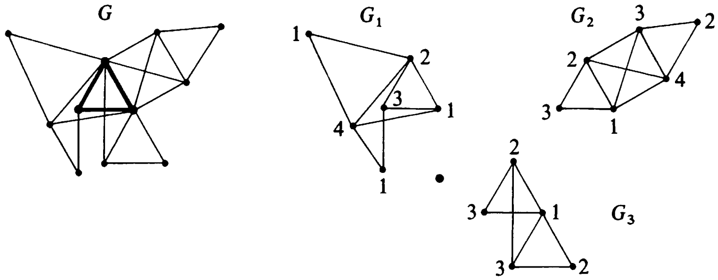


Figure V.2. The vertex set of the thick triangle disconnects G;

$\chi(G) = \max\{\chi(G_1), \chi(G_2), \chi(G_3)\}$.

As a rather crude consequence of Theorem 1 we see that $\chi(G) \leq \Delta + 1$, where $\Delta = \Delta(G)$ is the maximum degree of G, since $\max_{H \subset G} \delta(H) \leq \Delta(G)$. Furthermore, if G is connected and not Δ-regular then clearly $\max_{H \subset G} \delta(H) \leq \Delta - 1$ so $\chi(G) \leq \Delta$. The following result, due to Brooks, takes care of the regular case.

Theorem 3. *Let G be a connected graph with maximal degree Δ. Suppose G is neither a complete graph nor an odd cycle. Then $\chi(G) \leq \Delta$.*

PROOF. We know already that we may assume without loss of generality that G is 2-connected and Δ-regular. Furthermore, we may assume that $\Delta \geq 3$, since a 2-regular 3-chromatic graph is an odd cycle.

If G is 3-connected, let x_n be any vertex of G and let x_1, x_2 be two non-adjacent vertices in $\Gamma(x_n)$. If G is not 3-connected, let x_n be a vertex for which $G - x_n$ is separable, so has at least two blocks. Since G is 2-connected, each endblock of $G - x_n$ has a vertex adjacent to x_n. Let x_1 and x_2 be such vertices belonging to different endblocks.

In either case we have found vertices x_1, x_2 and x_n such that $G - \{x_1, x_2\}$ is connected, $x_1x_2 \notin E(G)$ but $x_1x_n \in E(G)$ and $x_2x_n \in E(G)$. Let $x_{n-1} \in V - \{x_1, x_2, x_n\}$ be a neighbour of x_n, let x_{n-2} be a neighbour of x_n or x_{n-1}, etc. Then the order $x_1, x_2, x_3, \ldots, x_n$ is such that each vertex other than x_n is adjacent to at least one vertex following it. Thus the greedy algorithm will use at most Δ colours, since x_1 and x_2 get the same colour and x_n is adjacent to both. □

Another colouring algorithm can be obtained by reducing the problem to colouring two other graphs derived from G. This reduction also enables us to obtain some information about the number of colourings of a graph with a given set of colours.

Let a and b be non-adjacent vertices of a graph G. Let G' be obtained from G by joining a to b, and let G'' be obtained from G by identifying a and b. Thus in G'' there is a new vertex ab instead of a and b, which is joined to the vertices adjacent to at least one of a and b (Figure V.3). The colourings of G in which a and b get *distinct* colours are in 1-1 correspondence with the colourings of G'. Indeed $c: V(G) \to \{1, 2, \ldots, k\}$ is a colouring of G with $c(a) \neq c(b)$ iff c is a colouring of G'. Similarly the colourings of G in which

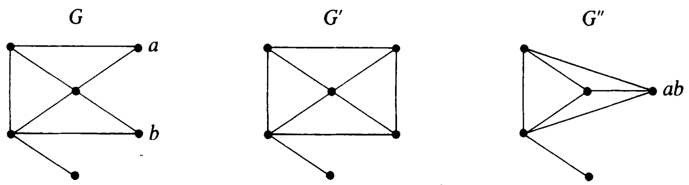


Figure V.3. The graphs G' and G''.

a and b get the *same* colour are in a 1-1 correspondence with the colourings of G''. In particular, if for a natural number x and a graph H we write $p_H(x)$ for the number of colourings of a graph H with colours $1, 2, \ldots, x$, then

$$p_G(x) = p_{G'}(x) + p_{G''}(x). \tag{1}$$

By definition $\chi(G)$ is the least natural number k for which $p_G(k) \geq 1$. Thus both the remarks above and relation (1) imply that

$$\chi(G) = \min\{\chi(G'), \chi(G'')\}. \tag{2}$$

The basic properties of $p_H(x)$ are given in our next result.

Theorem 4. *Let H be a graph with $n \geq 1$ vertices, m edges and k components. Then*

$$p_H(x) = \sum_{i=0}^{n-k} (-1)^i a_i x^{n-i},$$

where $a_0 = 1$, $a_1 = m$ and $a_i > 0$ for every i, $0 \leq i \leq n - k$.

PROOF. We apply induction on $n + m$. For $n + m = 1$ the assertions are trivial so we pass to the induction step. If $m = 0$, we are again home since in this case $k = n$ and as *every map* $f\colon V(H) \to \{1, 2, \ldots, x\}$ is a colouring of H, we have $p_H(x) = x^n$. If $m > 0$ we pick two adjacent vertices of H, say a and b. Putting $G = H - ab$ we find that $G' = H$. Since $e(G) = m - 1$ and $|G''| + e(G'') \leq n - 1 + m$, by the induction hypothesis the assertions of the theorem hold for $p_G(x)$ and $p_{G''}(x)$. Note now that G'' has k components and G has at least k components. Therefore

$$p_G(x) = x^n - (m - 1)x^{n-1} + \sum_{i=2}^{n-k} (-1)^i b_i x^{n-i}$$

where $b_i \geq 0$ for each i, and

$$p_{G''}(x) = x^{n-1} - \sum_{i=2}^{n-k} (-1)^i c_i x^{n-i},$$

where $c_i > 0$ for each i. Hence, by (1),

$$p_H(x) = p_{G'}(x) = p_G(x) - p_{G''}(x) = x^n - mx^{n-1} + \sum_{i=2}^{n-k} (-1)^i (b_i + c_i) x^{n-i}$$

$$= x^m - mx^{n-1} + \sum_{i=2}^{n-k} (-1)^i a_i x^{n-i},$$

where $a_i > 0$ for each i. □

Because of Theorem 4 we are justified in calling $p_H(x)$ the *chromatic polynomial* of H. The chromatic polynomial is closely connected to various enumeration problems, some of which will be discussed in Ch. VIII. In par-

ticular, the modulus of each coefficient of $p_H(x)$ is the number of certain subgraphs of H (cf. Exercise 12^+).

Let us describe now the algorithm based on the reduction $G \to \{G', G''\}$. Given a graph G, construct a sequence of graphs $G_0, G_1, \ldots$ as follows. Put $G_0 = G$. Having constructed G_i, if G_i is complete, terminate the sequence, otherwise let G_{i+1} be G_i' or G_i''. The sequence has to end in a complete graph G_t, say of order $|G_t| = k$. A k-colouring of G_t can easily be lifted to a k-colouring of the original graph G, so $\chi(G) \leq k$. In fact, equality (2) shows that $\chi(G)$ is the minimum order of a complete graph in which a sequence $G_0, G_1, \ldots$ can terminate.

There are other problems that can be tackled by the reduction $G \to \{G', G''\}$; a beautiful example is Exercise 16^+.

§2 Edge Colouring

In a colouring of the edges of a graph G the edges incident with a vertex get distinct colours, so $\chi'(G)$, the edge chromatic number, is at least as large as the maximum degree, $\Delta(G) = \max_x d(x)$,

$$\chi'(G) \geq \Delta(G). \tag{3}$$

At first sight it is somewhat surprising that this trivial inequality is in fact, an equality for large classes of graphs, including the class of *bipartite graphs*. Indeed, Exercise 22 of Chapter III, which is a consequence of Hall's theorem, asserts that the edge set $E(G)$ of a bipartite graph G can be partitioned into $\Delta(G)$ classes of independent edges, that is $\chi'(G) = \Delta(G)$.

Another trivial lower bound on $\chi'(G)$ follows from the fact that if G does not contain $\beta + 1$ independent edges, then each colour class has at most β edges so we need at least $\lceil e(G)/\beta \rceil$ colour classes to take care of all the edges:

$$\chi'(G) \geq \lceil e(G)/\beta \rceil. \tag{4}$$

Similarly to Theorem 9 of Chapter 1, it is easily shown that if G is a complete graph of order 2 then equality holds in (4), that is $\chi'(K^n) = n - 1$ if n is even and $\chi'(K^n) = n$ if $n \geq 3$ is odd (Exercise 28).

How can one obtain an upper bound for $\chi'(G)$? Since each edge is adjacent to at most $2(\Delta(G) - 1)$ edges, Theorem 1 implies that

$$\chi'(G) \leq 2\Delta(G) - 1.$$

Furthermore, if $\Delta(G) \geq 3$, the theorem of Broȯks gives

$$\chi'(G) \leq 2\Delta(G) - 2.$$

At first sight this inequality seems reasonably good. However, the following fundamental theorem of Vizing shows that this is not the case, for the edge-chromatic number is almost determined by the maximum degree.

Theorem 5. $\Delta(G) \le \chi'(G) \le \Delta(G) + 1$.

PROOF. Put $\Delta = \Delta(G)$ and assume that we have used $1, 2, \dots, \Delta + 1$ to colour all but one of the edges. We are home if we can show that this uncoloured edge can also be coloured.

We say that a colour is *missing at* a vertex z if no edge incident with z gets that colour. If z is incident with $d'(z) \le d(z) \le \Delta$ edges that have been coloured, then $\Delta + 1 - d'(z)$ colours are missing at z. In particular, at each vertex at least one colour is missing. Our aim is to move around the colours and the uncoloured edge in such a way that a colour will be missing at both endvertices of the uncoloured edge, enabling us to complete the colouring.

Let xy_1 be the uncoloured edge; let s be a colour missing at x and let t_1 be a colour missing at y_1. We shall construct a sequence of edges $xy_1, xy_2, \dots$, and a sequence of colours $t_1, t_2, \dots$ such that t_i is missing at y_i and xy_{i+1} has colour t_i. Suppose we have constructed $xy_1, \dots, xy_i$ and $t_1, \dots, t_i$. There is at most one edge xy of colour t_i. If $y \notin \{y_1, \dots, y_i\}$, we put $y_{i+1} = y$ and pick a colour t_{i+1} missing at y_{i+1}, otherwise we stop the sequence. These sequences have to terminate at most $\Delta(G)$ terms; let $xy_1, \dots, xy_h$ and $t_1, \dots, t_h$ be the complete sequences. Let us examine the two reasons that may have forced us to terminate these sequences.

(a) *No edge xy has colour t_h.* Then recolour the edges xy_i, $i < h$, giving xy_i the colour t_i. In the colouring we obtain every edge is coloured, except xy_h. However, since t_h occurs neither at x nor at y_h, we may complete the colouring by assigning t_h to xy_h.

(b) *For some $j < h$ the edge xy_j has colour t_h.* To start with, recolour the edges xy_i, $i < j$, giving xy_i the colour t_i. In this colouring the uncoloured edge is xy_j. Let $H(s, t_h)$ be the subgraph of G formed by the edges of colour s and t_h. Each vertex of $H(s, t_h)$ is incident with at most 2 edges in $H(s, t_h)$ (one of colour s and the other of colour t_h) so the components of $H(s, t_h)$ are paths and cycles. Each of the vertices x, y_j and y_h has degree at most 1 in $H(s, t_h)$, so they cannot belong to the same component of $H(s, t_h)$. Thus at least one of the following two cases has to hold.

(b1) *The vertices x and y_j belong to distinct components of* $H(s, t_h)$. Interchange the colours s and t_h in the component containing y_j. Then the colour s is missing at both x and y_j, so we may complete the colouring by giving xy_j the colour s.

(b2) *The vertices x and y_h belong to distinct components* of $H(s, t_h)$. Continue the recolouring of the edges incident with x by giving xy_i the colour t_i for each $i < h$, thereby making xy_h the uncoloured edge. This change does not involve edges of colour s and t_h, so $H(s, t_h)$ has not been altered. Now switch around the colours in the component containing y_h. This switch makes sure that the colour s is missing at both x and y_h, so we can use s to colour the so far uncoloured edge xy_h. □

Note that the proof above gives an algorithm for colouring the edges with at most $\Delta + 1$ colours.

§3 Graphs on Surfaces

The most famous problem in graph theory is undoubtedly the four colour problem: prove that every plane graph is 4-colourable. The weaker assertion, namely that every plane graph is 5-colourable, is an almost immediate consequence of Euler's formula.

Theorem 6. *Every plane graph is 5-colourable.*

PROOF. Suppose the assertion is false and let G be a 6-chromatic plane graph with minimal number of vertices. By Theorem 12 of Chapter I G has a vertex x of degree at most 5. Put $H = G - x$. Then H is 5-colourable, say with colours $1, 2, \ldots, 5$. Each of these colours must be used to colour at least one neighbourhood of x, otherwise the missing colour could be used to colour x. Hence we may assume that x has 5 neighbours, say $x_1, x_2, \ldots, x_5$ in some cyclic order about x, and the colour of x_i is i, $i = 1, 2, \ldots, 5$. Denote by $H(i, j)$ the subgraph of H spanned by the vertices of colour i and j.

Suppose first that x_1 and x_3 belong to distinct components of $H(1, 3)$. Interchanging the colours 1 and 3 in the component of x_1, we obtain another 5-colouring of H. However, in this 5-colouring both x_1 and x_3 get colour 3, so 1 is not used to colour any of the vertices $x_1, \ldots, x_5$. This is impossible since then x can be coloured 1.

Since x_1 and x_3 belong to the same component of $H(1, 3)$, there is an x_1-x_3 path P_{13} in H whose vertices are coloured 1 and 3. Analogously H contains an x_2-x_4 path P_{24} whose vertices are coloured 2 and 4. However, this is impossible, since the cycle $xx_1P_{13}x_3$ of G separates x_2 from x_4 but P_{24} cannot meet this cycle (Figure V.4). □

Clearly not every plane graph is 3-colourable. Indeed, K^4 is planar and it does need 4 colours. Another 4-chromatic planar graph is obtained by joining all five vertices of a C^5 to a sixth vertex. Thus if we want to determine $\chi_0 = \max\{\chi(G): G \text{ is planar}\}$ then we see immediately that $\chi_0 \geq 4$ and $\chi_0 \leq 5$ and the problem is to prove $\chi_0 \leq 4$.

Instead of a plane graph we may wish to consider a graph drawn on an orientable surface of genus $\gamma > 0$ and ask for χ_γ, the *maximum* of the chromatic

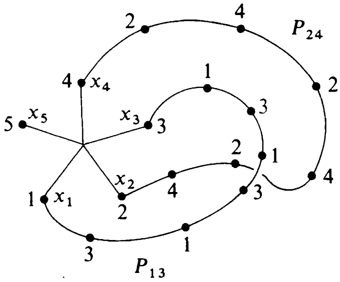


Figure V.4. The paths P_{13} and P_{24}.

number of such a graph. We shall see in a moment that in this problem we encounter a rather different difficulty. For our limited purpose it will suffice to remark that an *orientable surface* S_γ of *genus* $\gamma \geq 1$ is obtained from a sphere by attaching to it γ "handles", and a graph of order n drawn on S_γ has at most $3n + 6(\gamma - 1)$ edges. Thus S_1 is the torus, and every toroidal graph (that is, graph that can be drawn on a torus) of order n has at most $3n$ edges. The following easy upper bound on the chromatic number of a graph drawn on an orientable surface of genus $\gamma \geq 1$ was obtained by Heawood in 1890.

Theorem 7. *The number of colours needed to colour a graph drawn on an orientable surface of genus $\gamma \geq 1$ is at most*

$$H(\gamma) = \lfloor \tfrac{1}{2}\{7 + \sqrt{1 + 48\gamma}\} \rfloor.$$

PROOF. Let G be the graph in question and let H be any subgraph of G. By Theorem 1 we are home if we show that $\delta(H) \leq H(\gamma) - 1$. If H has n vertices, it has at most $3n + 6(\gamma - 1)$ edges so

$$\delta(H) \leq \min\left\{n - 1, 6 + \left\lfloor \frac{12(\gamma - 1)}{n} \right\rfloor\right\}.$$

Hence the result follows if

$$6 + \left\lfloor \frac{12(\gamma - 1)}{H(\gamma) + 1} \right\rfloor \leq H(\gamma) - 1.$$

With a little work one can check that this inequality indeed holds. □

By applying a stronger colouring result than the trivial Theorem 1, one can show that $H(\gamma) - 1$ colours suffice to colour a graph on S_γ *unless* the graph contains a complete graph of order $H(\gamma)$. Thus $\chi_\gamma = H(\gamma)$ iff $K^{H(\gamma)}$ can be drawn on S_γ. Heawood gave a false proof of the assertion $\chi_\gamma = H(\gamma)$; the first correct proof, due to Ringel and Youngs, was found only over 75 years later. Note that the difficulty in proving this deep result lies in finding a drawing of a fixed and single graph, $K^{H(\gamma)}$, on a surface of genus $\gamma > 0$. On the other hand, in order to solve the four colour problem one has to show that *every* plane graph can be coloured with four colours. Thus the problem of determining χ_0 has almost nothing to do with the problem of determining χ_γ for $\gamma \geq 1$.

Before saying a few words about the solution of the four colour problem, let us point out that $H(1) = 7$ is easily proved; in fact, this was indeed proved by Heawood. All one needs is that K^7 can be drawn on the torus. Recalling that a torus can be obtained from a rectangle by identifying opposite sides, Figure V.5 shows a required drawing.

We saw in Chapter I. §4 that a plane graph G determines a *map* $M = M(G)$ consisting of the plane graph G and the countries determined by the plane graph. A colouring of a map is a colouring of the countries such that no two

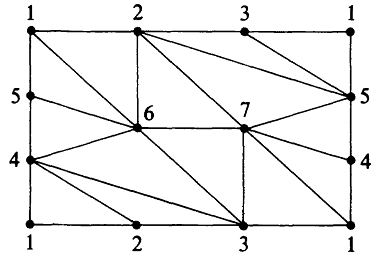


Figure V.5. A drawing of K^7 on the torus.

countries sharing an edge in their boundaries get the same colour. Show that every plane map can be coloured with four colours. This is the original form of the four colour problem, as posed by Francis Guthrie in 1852. His teacher, de Morgan, circulated it amongst his colleagues, but the problem was first made popular in 1878 by Cayley, who mentioned it before the Royal Society. Almost at once "proofs" appeared, by Kempe in 1879 and by Tait in 1880. Heawood's refutation of Kempe's proof was published in 1890, though he modified the proof to obtain the five colour theorem. Tait's paper also contained false assumptions which prompted Petersen to observe in 1891 that the four colour theorem is equivalent to the conjecture that every cubic planar graph has edge chromatic number three (Exercise 27). Contributions to the solution since the turn of the century include Birkhoff's introduction of the chromatic polynomial and works by various authors giving lower bounds on the order of a possible counterexample. In 1943 Hadwiger made a deep conjecture containing the four colour problem as a special case: if $\chi(G) = k$ then G is contractible to K^k (see Exercises 18–20).

In hindsight the most important advance was made by Heesch, who developed the method of discharging to find unavoidable sets of reducible configurations. The problem was at last solved by Appel and Haken in 1976, making use of a refinement of Heesch's method and fast electronic computers. The interested reader is referred to some papers of Appel and Haken, and to the book of Saaty and Kainen for a detailed explanation of the underlying ideas of the proof. All we have room for is a few superficial remarks.

What makes the five colour theorem true? The following two facts: (i) a minimal 6-colourable plane graph cannot contain a vertex of degree at most 5, and (ii) a plane graph has to contain a vertex of degree at most 5. We can go a step further and ask why (i) and (ii) hold. A look at the proof shows that (i) is proved by making a good use of the paths P_{ij}, called *Kempe chains* after Kempe, who used them in his false proof of 1879, and (ii) follows immediately from Euler's formula $n - e + f = 2$.

The attack on the four colour problem goes along similar lines. A *configuration* is a connected cluster of vertices of a plane graph together with the degrees of the vertices. A configuration is *reducible* if no minimal 5-chromatic plane graph can contain it, and a set of configurations is *unavoidable* if every plane graph contains at least configuration belonging to

the set. In order to prove that every plane graph is 4-colourable, one sets out to find an *unavoidable set of reducible configurations*. How should one show that a configuration is reducible? Replace the cluster of vertices by a smaller cluster, 4-colour the obtained smaller graph and use Kempe chains to show that the 4-colouring can be "pulled back" to the original graph. How should one show that a set of configurations is unavoidable? Make extensive use of Euler's formula. Of course, one may always assume that the graph is a maximal plane graph. Assigning a *charge* of 6-k to a vertex of degree k, Euler's formula guarantees that the total charge is 12. Push charges around the vertices, that is transfer some charge from a vertex to some of its neighbours, until it transpires that the plane graph has to contain one of the configurations.

Looking again at the five colour theorem, we see that the proof was based on the fact that set of configurations of single vertices of degree at most 5 is an unavoidable set of configurations (for the five colour theorem).

The simplistic sketch above does not indicate the difficulty of the actual proof. In order to rectify this a little, we mention that Appel and Haken had to find over 1900 reducible configurations to complete the proof. Furthermore, we invite the reader to prove the following two simple assertions.

1. The configurations in Figure V.6 are reducible.

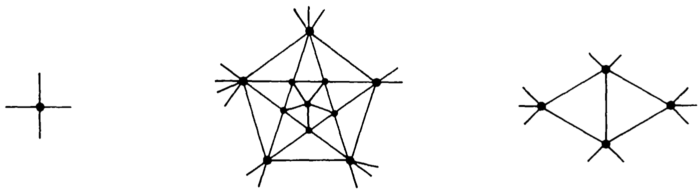


Figure V.6. Three reducible configurations; in the second examples the outer vertices may have arbitrary degrees.

2. Let G be a maximal planar graph of order at most 25 and minimum degree 5. Call a vertex a *major* vertex if its degree is at least 7, otherwise call it a *minor*. Then G contains one of the following:

   (a) a minor vertex with 3 consecutive neighbours of degree 5,
   (b) a vertex of degree 5 with minor neighbours only,
   (c) a major vertex with at most one neighbour of degree at least 6.

EXERCISES

1.⁻ Show that a graph G has at least $\binom{\chi(G)}{2}$ edges.

2. For each $k \geq 3$ find a bipartite graph with vertices $x_1, x_2, \ldots, x_n$ for which the greedy algorithm uses k colours. Can it be done with $n = 2k - 2$? Show that it cannot be done with $n = 2k - 3$.

3.$^-$ Given a graph G, order its vertices in such a way that the greedy algorithm uses only $k = \chi(G)$ colours.

4.$^-$ Let $d_1 \geq d_2 \geq \cdots \geq d_n$ be the degree sequence of G. Show that in an order $x_1, x_2, \ldots, x_n$, $d(x_i) = d_i$, the greedy algorithm uses at most $\max \min\{d_i + 1, i\}$ colours, and so if k is the maximal natural number for which $k \leq d_k + 1$ then $\chi(G) \leq k$.

5. Deduce from Exercise 4 that if G has n vertices then

$$\chi(G) + \chi(\bar{G}) \leq n + 1.$$

6.$^-$ Show that $\chi(G) + \chi(\bar{G}) \geq 2\sqrt{n}$.

7. Let d, d_1 and d_2 be non-negative integers with $d_1 + d_2 = d - 1$. Prove that if $\Delta(G) = d$ then the vertex set $V(G)$ of G can be partitioned into two classes, say $V(G) = V_1 \cup V_2$, such that the graphs $G_i = G[V_i]$ satisfy $\Delta(G_i) \leq d_i$, $i = 1, 2$. [*Hint*. Consider a partition $V(G) = V_1 \cup V_2$ for which $d_1 e(G_2) + d_2 e(G_1)$ is minimal.]

8. (Exercise 7 continued) Let now $d, d_1, d_2, \ldots, d_r$ be non-negative integers with $\sum_1^r (d_i + 1) = d + 1$. Prove that if $\Delta(G) = d$ then there is a partition $V(G) = \bigcup_1^r V_i$ such that the graphs $G_i = G[V_i]$ satisfy $\Delta(G_i) \leq d_i$, $i = 1, 2, \ldots, r$.

9. Given natural numbers r and t, $2r \leq t$, the *Kneser graph* $K_t^{(r)}$ is constructed as follows. Its vertex set is $T^{(r)}$, the set of r-element subsets of $T = \{1, 2, \ldots, t\}$, and two vertices are joined iff they are disjoint subsets of T. Figure V.7 shows $K_5^{(2)}$, the so called *Petersen graph*. Prove that $\chi(K_t^{(r)}) \leq t - 2r + 2$, $\chi(K_5^{(2)}) = 3$ and $\chi(K_6^{(2)}) = 4$.

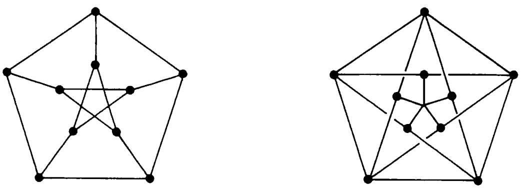


Figure V.7. The Petersen graph and the Grötzsch graph.

10. Check that the Grötzsch graph (Figure V.7) has girth 4 and chromatic number 4. Show that there is no graph of order 10 with girth at least 4 and chromatic number 4.

11.$^+$ Try to construct a triangle free graph of chromatic number 1526 without looking at Chapters VI or VII.

12.$^+$ Let H be a graph of order n with edges $e_1, e_2, \ldots, e_m$ and chromatic polynomial $p_H(x) = \sum_{i=0}^n (-1)^i c_i x^{n-i}$. Call a set of edges a *broken cycle* if it is obtained from the edge set of a cycle by omitting the edge of highest index. Show that c_j is the number of j-element subsets of the edges containing no broken cycle [*Hint*. Imitate the proof of Theorem 4, choosing e_1 for ab. Put $G = H - e_1$ so that $H = G'$ and $p_H(x) = p_G(x) - p_{G''}(x)$. If an edge of G'' comes from two edges of G, say from $xa = e_i$ and $xb = e_h$, label it with e_k, where $k = \max\{i, h\}$; otherwise keep the label it had in G (and H). Let F be a set of edges of H containing no

broken cycle of H. Note that if $e_1 \notin F$ then F is a set of edges of G containing no broken cycle, and if $e_1 \in F$ then $F - \{e_1\}$ is a set of edges of G'' containing no broken cycle.]

13. Let $p_G(x) = \sum_{i=0}^{n} (-1)^i a_i x^{n-i}$ be the chromatic polynomial of G. Prove that $a_2 = \binom{m}{2} - k_3(G)$, where $m = e(G)$ and $k_3(G)$ is the number of triangles in G.

14. Show that there is a unique graph G_0 of order n and size $m = \lfloor n^2/4 \rfloor$ such that if G is also of order n and size m then

$$p_G(x) \leq p_{G_0}(x)$$

whenever x is sufficiently large.

15.$^-$ Find graphs G and H of order n and same size such that $\chi(G) < \chi(H)$ but $p_G(x) < p_H(x)$ if x is sufficiently large.

16.$^+$ Given a connected graph G containing at least one cycle, define a graph H on the set S of all spanning trees of G by joining T_1 to T_2 iff $|E(T_1)\backslash E(T_2)| = 1$. (Cf. simple transform of an x-path, p. 79). Imitate the proof of the fact that $p_H(x)$ is a polynomial (Theorem 4) and the proof suggested in Exercise 12^+, to show that H is not only Hamiltonian, but every edge of it is contained in a Hamiltonian cycle.

17.$^+$ We say that a graph G has a subgraph *contractible* to a graph H with vertex set $\{y_1, \ldots, y_k\}$ if G contains vertex disjoint connected subgraphs $G_1, \ldots, G_k$ such that, for $i \neq j$, $y_i y_j \in E(H)$ iff G has a G_i-G_j edge.

Prove that for every natural number p there is a minimal integer $c(p)$ such that every graph with chromatic number at least $c(p)$ has a subgraph contractible to K^p. Show that $c(1) = 1$, $c(2) = 2$ and $c(n+1) \leq 2c(n) - 1$ for $n \geq 2$. [*Hint*. Pick $x_0 \in G$ and let S_k be the "sphere" of centre x_0 and radius k; $S_k = \{x \in G: d(x_0, x) = k\}$. Show that $\chi(G) \leq \chi(S_k) + \chi(S_{k+1})$ for some k.]

18.$^+$ *Hadwiger's conjecture* states that $c(p) = p$ for every p. Prove this for $p \leq 4$.

19. Deduce from Kuratowski's theorem, Theorem 13 of Chapter I, that a graph is planar iff it has no subgraph contractible to K^5 or $K^{3,3}$.

20. Show that the truth of Hadwiger's conjecture for $p = 5$ implies the four colour theorem.

21.$^-$ Show that a map $M = M(G)$ can be 2-coloured (see p. 96) iff every vertex of G has even degree.

22.$^+$ Suppose G is a cubic plane graph. Prove that the map $M(G)$ is 3-colourable iff each country has an even number of sides.

23.$^-$ Let $M = M(G)$ be a triangular map, that is a map in which every country has three sides. Show that M is 3-colourable unless $G = K^4$.

24.$^-$ Prove that a map $M = M(G)$ is 4-colourable if G has a Hamilton cycle.

25.$^-$ For each plane graph G construct a cubic plane graph H such that if $M(H)$ is is 4-colourable then so is $M(G)$.

26.⁻ According to Tait's conjecture every 3-connected cubic plane graph has a Hamilton cycle. (i) Show that Tait's conjecture implies the four colour theorem. (ii) By examining the graph in Figure V.8 disprove Tait's conjecture.

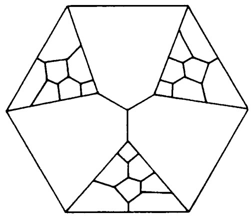


Figure V.8. Tutte's counterexample to Tait's conjecture

27.⁺ Let G be a cubic plane graph. Show that G is 3-edge-colourable iff $M(G)$ is 4-colourable. [*Hint.* Let $1, a, b$ and c be the elements of the Klein fourgroup $C_2 \times C_2$, so that $a^2 = b^2 = c^2 = 1$. Colour the edges with a, b and c, and the ountries with $1, a, b$ and c.]

28.⁻ Find the edge chromatic number of K^n.

29. The cubic graph G has exactly one edge colouring with $\chi'(G)$ colours. Show that $\chi'(G) = 3$ and that G has exactly 3 Hamilton cycles.

30. Let $n = 2^p$. Show that K^{n+1} is not the union of p bipartite graphs but K^n is. Deduce that if there are $2^p + 1$ points in the plane then some three of them determine an angle of size at least $\pi(1 - (1/p))$.

31. Let $\chi(G) = k$. What is the minimal number of r-chromatic graphs whose union is G?

32. Show that a k-chromatic graph can be oriented in such a way that a longest directed path has k vertices.

33. Show that if a graph G can be oriented in such a way that no directed path contains more than k vertices than $\chi(G) \leq k$. [*Hint.* Omit a minimal set of edges to destroy all directed cycles. For a vertex x let $c(x)$ be the maximal number of vertices on a directed path in the new graph starting at x. Check that c is a proper colouring.]

34. Denote by $\mathscr{G}_k$ the set of (isomorphism classes of) graphs of chromatic number at least k. Let G_1 and G_2 be vertex disjoint graphs in $\mathscr{G}_k$ and let $a_i b_i$ be an edge of G_i, $i = 1, 2$. Let $G = G_1 \nabla G_2$ be obtained from $G_1 \cup G_2$ by omitting the edges

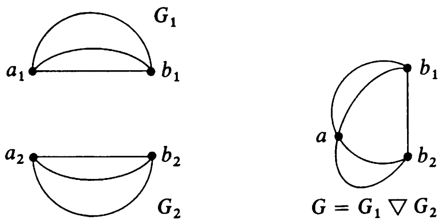


Figure V.9. The Hajós operation.

a_1b_1, a_2b_2, adding the edge b_1b_2 and identifying the vertices a_1 and a_2. (The operation ∇ is said to be the *Hajós operation*, see Figure V.9.) Prove that $G \in \mathscr{G}_k$.

35.⁺ Let $\mathscr{H}_k$ be the smallest collection of (isomorphism classes of) graphs such that (1) $K^k \in \mathscr{H}_k$, (2) if $H \in \mathscr{H}_k$ and $G \supset H$ then $G \in \mathscr{H}_k$, (3) if $H \in \mathscr{H}_k$ and G is obtained from H by identifying two non-adjacent vertices, then $G \in \mathscr{H}_k$, (4) if $G_1, G_2 \in \mathscr{H}_k$ and $G = G_1 \nabla G_2$ then $G \in \mathscr{H}_k$. Prove that $\mathscr{H}_k = \mathscr{G}_k$. [*Hint.* Exercise 33 implies that $\mathscr{H}_k \subset \mathscr{G}_k$. Assume that the converse inclusion is false and let $G \in \mathscr{G}_k \backslash \mathscr{H}_k$ be a counterexample of minimal order and maximal size. G cannot be a complete q-partite graph so it contains vertices a, b_1 and b_2 such that $b_1b_2 \in E(G)$ but $ab_1, ab_2 \notin E(G)$. Let $G_1 = G + ab_1$ and $G_2 = G + ab_2$. Then G_1 and G_2 are not counterexamples so belong to $\mathscr{H}_k$. Find out how G can be obtained from copies of G_1 and G_2 by the allowed operations.]

36. Let k be a natural number. Prove that an infinite graph is k-colourable iff every finite subgraph of it is. [*Hint.* Apply Tychonov's theorem as in Exercise 24 of Chapter III.]

Notes

Theorem 3 is in R. L. Brooks, On colouring the nodes of a network, *Proc. Cambridge Phil. Soc.* **37** (1941) 194–197, and Vizing's theorem, Theorem 5, is in V. G. Vizing, On an estimate of the chromatic class of a p-graph (in Russian) *Diskret. Analiz* **3** (1964) 23–30. A detailed account of results concerning colouring graphs on surfaces, culminating in Ringel and Youngs' proof of Heawood's conjecture, can be found in *Map Color Theorem*, Grundlehren der math. Wiss. 209, Springer-Verlag, Berlin, 1974. The use of the discharging procedure in attacking the four colour problem is described in H. Heesch, Untersuchungen zum Vierfarbenproblem, B-I-Hochschulskripten 810/810a/810b, Bibliographisches Institut, Mannheim, Vienna, Zürich, 1969. The proof of the four colour theorem appears in K. Appel and W. Haken, Every planar map is four colourable, Part I: discharging, *Illinois J. of Math.* **21** (1977) 429–490 and K. Appel, W. Haken and J. Koch, Every planar map is four colourable, Part II: reducibility, *Illinois J. of Math.* **21** (1977) 491–567. A history of the four colour problem and a digest of its proof are provided by T. L. Saaty and P. C. Kainen, *The Four Color Problem, Assaults and Conquest*, McGraw-Hill, New York, 1977. The partition results of Exercises 7 and 8 are in L. Lovász, *On Decomposition of Graphs*, Studia Sci. Math., Hungar. **1** (1966) 237–238.

The result in Exercise 33 is due to T. Gallai, On directed paths and circuits, in *Theory of Graphs* (P. Erdös and G. Katona, eds.), Academic Press, New York, 1968, 115–118, and that of Exercise 35 was proved by G. Hajós, Über eine Konstruktion nicht n-färbbarer Graphen, *Wiss. Zeitschr.* Martin Luther Univ. Halle-Wittenberg, Math.-Natur. Reihe **10** (1961) 116–117.

Colouring is a naturally appealing part of graph theory and the subject has a vast literature. Many of the fundamental results are due to G. A. Dirac; for these and other results see Chapter V. of B. Bollobás, *Extremal Graph Theory*, Academic Press, London and New York, 1978.

CHAPTER VI
Ramsey Theory

Show that in a party of six people there is always a group of three who either all know each other or are all strangers to each other. This well known puzzle is a special case of a theorem proved by Ramsey in 1928. The theorem has many deep extensions which are important not only in graph theory and combinatorics but in set theory (logic) and analysis as well. In this chapter we prove the original theorems of Ramsey, indicate some variations and present some applications of the results.

§1 The Fundamental Ramsey Theorems

We shall consider partitions of the *edges* of graphs and hypergraphs. For the sake of convenience a partition will be called a *colouring*, but one should bear in mind that a colouring in this sense has nothing to do with the edge colourings considered in Chapter V. Adjacent edges may have the same colour and, indeed, our aim is to show that there are large subgraphs all of whose edges have the same colour. In a 2-colouring we shall always choose red and blue as colours; a *subgraph* is *red* (*blue*) if all its edges are red (blue).

Given a natural number s, there is an $n(s)$ such that if $n \geq n(s)$ then every colouring of the edges of K^n with red and blue contains either a red K^s or a blue K^s. In order to show this and to give a bound on $n(s)$, we introduce the following notation: $R(s, t)$ is the minimum of n for which every red-blue colouring of K^n yields a red K^s or a blue K^t. (We assume that $s, t \geq 2$, for we adopt the reasonable convention that every K^1 is both red and blue since it has no edges.) A priori it is not clear that $R(s, t)$ is finite for every s and t.

However, it is obvious that

$$R(s, t) = R(t, s) \quad \text{for every } s, t \geq 2$$

and

$$R(s, s) = R(2, s) = s,$$

since in a red-blue colouring of K^s either there is a blue edge or else every edge is red. The following result shows that $R(s, t)$ is finite for every s and t, and at the same time it gives a bound on $R(s, t)$.

Theorem 1. *If $s > 2$ and $t > 2$ then*

$$R(s, t) \leq R(s - 1, t) + R(s, t - 1) \tag{1}$$

and

$$R(s, t) \leq \binom{s + t - 2}{s - 1}. \tag{2}$$

Proof. (i) When proving (1) we may assume that $R(s - 1, t)$ and $R(s, t - 1)$ are finite. Let $n = R(s - 1, t) + R(s, t - 1)$ and consider a colouring of the edges of K^n with red and blue. Wd have to show that this colouring contains either a red K^s or a blue K^t. Let x be a vertex of K^n. Since $d(x) = n - 1 = R(s - 1, t) + R(s, t - 1) - 1$, either there are at least $n_1 = R(s - 1, t)$ red edges incident with x or there are at least $n_2 = R(s, t - 1)$ blue edges incident with x. By symmetry we may assume that the first case holds. Consider a subgraph K^{n_1} of K^n spanned by n_1 vertices joined to x by a red edge. If K^{n_1} has a blue K^t, we are home. Otherwise K^{n_1} contains a red K^{s-1} which forms a red K^s with x.

(ii) Inequality (2) holds if $s = 2$ ot $t = 2$ (in fact, we have equality since $R(s, 2) = R(2, s) = s$). Assume now that $s > 2$, $t > 2$ and (2) holds for every pair (s', t') with $2 \leq s', t'$ and $s' + t' < s + t$. Then by (1) we have

$$\begin{aligned} R(s, t) &\leq R(s - 1, t) + R(s, t - 1) \\ &\leq \binom{s + t - 3}{s - 2} + \binom{s + t - 3}{s - 1} = \binom{s + t - 2}{s - 1}. \end{aligned} \qquad \square$$

The result easily extends to colourings with arbitrarily (but finitely) many colours: given k and $s_1, s_2, \ldots, s_k$, if n is sufficiently large then every colour of K^n with k colours is such that for some i, $1 \leq i \leq k$, it contains a K^{s_i} coloured with the i-th colour. (The minimal value of n for which this holds is usually denoted by $R_k(s_1, \ldots, s_k)$.) Indeed, if we know this for $k - 1$ colours, then in a k-colouring of K^n we replace the first two colours by a new colour. If n is sufficiently large (depending on $s_1, s_2, \ldots, s_k$) then either there is a K^{s_i} coloured with the i-th colour for some i, $3 \leq i \leq k$, or else there is a

K^m, $m = R(s_1, s_2)$ coloured with the new colour, that is coloured with the first two (original) colours. In the first case we are home and in the second, for $i = 1$ or 2 we can find a K^{s_i} in K^m coloured with the i-th colour.

In fact, Theorem 1 also extends to hypergraphs, that is to colourings of the set $X^{(r)}$ *of all r-tuples* of a finite set X with k colours. This is one of the theorems proved by Ramsey.

Denote by $R^{(r)}(s, t)$ the minimum of n for which every red-blue colouring of $X^{(r)}$ yields a red s-set or a blue t-set, provided $|X| = n$. Of course, a set $Y \subset X$ is called red (blue) if every element of $Y^{(r)}$ is red (blue). Note that $R(s, t) = R^{(2)}(s, t)$. As in the case of Theorem 1, the next result not only guarantees that $R^{(r)}(s, t)$ is finite for all values of the parameters (which is not at all obvious a priori), but also gives an upper bound on $R^{(r)}(s, t)$. The proof is an almost exact replica of the proof of Theorem 1. Note that if $r > \min\{s, t\}$ then $R^{(r)}(s, t) = \min\{s, t\}$ and if $r = s \leq t$ then $R^{(r)}(s, t) = t$.

Theorem 2. *Let $1 < r < \min\{s, t\}$. Then $R^{(r)}(s, t)$ is finite and*

$$R^{(r)}(s, t) \leq R^{(r-1)}(R^{(r)}(s - 1, t), R^{(r)}(s, t - 1)) + 1.$$

Proof. Both assertions follow if we prove the inequality under the assumption that $R^{(r-1)}(u, v)$ is finite for all u, v, and $R^{(r)}(s - 1, t)$, $R^{(r)}(s, t - 1)$ are also finite.

Let X be a set with $R^{(r-1)}(R^{(r)}(s - 1, t), R^{(r)}(s, t - 1)) + 1$ elements. Given any red-blue colouring of $X^{(r)}$, pick an $x \in X$ and define a red-blue colouring of the $(r - 1)$-sets of $Y = X - \{x\}$ by colouring $\sigma \in Y^{(r-1)}$ the colour of $\{x\} \cup \sigma \in X^{(r)}$. By the definition of the function $R^{(r-1)}(u, v)$ we may assume that Y has a red subset Z with $R^{(r)}(s - 1, t)$ elements.

Now let us look at the colouring of $Z^{(r)}$. If it has a blue t-set, we are home, since $Z^{(r)} \subset X^{(r)}$ so a blue t-set of Z is also a blue t-set of X. On the other hand if there is no blue t-set of Z then there is a red $(s - 1)$-set, and its union with $\{x\}$ is then a red s-set of X. □

Very few of the non-trivial Ramsey numbers are known, even in the case $r = 2$. It is easily seen that $R(3, 3) = 6$ and with some work one can show that $R(3, 4) = 9$, $R(3, 5) = 14$, $R(3, 6) = 18$, $R(3, 7) = 23$ and $R(4, 4) = 18$. Because of (1) any upper bound on an $R(s, t)$ helps to give an upper bound on every $R(s', t')$, $s' \geq s$, $t' \geq t$. Lower bounds for $R(s, t)$ are not easy to come by either. In Chapter VII we shall apply the method of random graphs to obtain lower bounds on the Ramsey numbers $R(s, t)$.

As a consequence of Theorem 2 we see that in every red-blue colouring of the r-tuples of the natural numbers contains arbitrarily large monochromatic subsets; a subset if *monochromatic* if its r-tuples have the same colour. Ramsey proved that, in fact, we can find an *infinite monochromatic* set.

Theorem 3. *Let $c: A^{(r)} \to \{1, \ldots, k\}$ be a k-colouring of the r-tuples $(1 \leq r < \infty)$ of an infinite set A. Then A contains a monochromatic infinite set.*

PROOF. We apply induction on r. Note that the result is trivial for $r = 1$ so assume that $r > 1$ and the theorem holds for smaller values of r.

Put $A_0 = A$ and pick an element $x_1 \in A_0$. As in the proof of Theorem 2, define a colouring $c_1: B_1^{(r-1)} \to \{1, \ldots, k\}$ of the $(r-1)$-tuples of $B_1 = A_0 - \{x_1\}$ by putting $c_1(\tau) = c(\tau \cup \{x_1\})$, $\tau \in B_1^{(r-1)}$. By the induction hypothesis B_1 contains an infinite set A_1 all whose $(r-1)$-tuples have the same colour, say d_1 ($d_1 \in \{1, \ldots, k\}$). Let now $x_2 \in A_1$, $B_1 = A_1 - \{x_2\}$ and define a k-colouring $c_2: B_2^{(r-1)} \to \{1, \ldots, k\}$ by putting $c_2(\tau) = c(\tau \cup \{x_2\})$, $\tau \in B_2^{(r-1)}$. Then B_2 has an infinite set A_2 all of whose $(r-1)$-tuples have the same colour, say d_2. Continuing in this way we obtain an infinite sequence of elements: $x_1, x_2, \ldots$, an infinite sequence of colours: $d_1, d_2, \ldots$, and an infinite nested sequence of sets: $A_0 \supset A_1 \supset A_2 \supset \cdots$, such that $x_i \in A_{i-1}$ and all r-tuples whose only element outside A_u is x_i have the same colour d_i. The infinite sequence $(d_n)_1^\infty$ must take at least one of the k values $1, 2, \ldots, k$ infinitely often, say $n_i \to \infty$ and $d_{n_i} = 1$ for every i. Then by the construction each r-tuple of the infinite set $\{x_{n_1}, x_{n_2}, \ldots\}$ has colour 1. □

In some cases it is more convenient to apply the following version of Theorem 3. ($\mathbb{N}$ is the set of natural numbers.)

Theorem 3′. *Let $k_r \in \mathbb{N}$, $r \in \mathbb{N}$, and colour the set $\mathbb{N}^{(r)}$ of r-tuples of $\mathbb{N}$ with k_r colours. Then there is an infinite set $M \subset \mathbb{N}$ such that for every r any two r-tuples of M have the same colour, provided their minimal elements are at least r.*

PROOF. Put $M_0 = \mathbb{N}$. Having chosen infinite sets $M_0 \supset \cdots \supset M_{r-1}$, let M_r be an infinite subset of M_{r-1} such that all the r-tuples of M_r have the same colour. This way we obtain an infinite nested sequence of infinite sets: $M_0 \supset M_1 \supset \cdots$. Pick $a_1 \in M_1, a_2 \in M_2 - \{1, \ldots, a_1\}, a_3 \in M_3 - \{1, \ldots, a_2\}$, etc. Clearly $M = \{a_1, a_2, \ldots\}$ has the required properties. □

Either Theorem 2 and the colour-grouping argument described after Theorem 1, or Theorem 3 and a compactness argument imply the following. Given r and $s_1, s_2, \ldots, s_k$, then for large enough $|X|$ every colouring of $X^{(r)}$ with k colours is such that for some i, $1 \le i \le k$, there is a set $S_i \subset X$, $|S_i| = s_i$, all of whose r-sets have colour i. The smallest value of $|X|$ for which this is true is denoted by $R_k^{(r)}(s_1, s_2, \ldots, s_k)$; thus $R^{(r)}(s, t) = R_2^{(r)}(s, t)$ and $R_k(s_1, s_2, \ldots, s_k) = R_k^{(2)}(s_1, s_2, \ldots, s_k)$. The upper bound for $R_k^{(r)}(s_1, s_2, \ldots, s_k)$ implied by Theorem 2 is not very good. Imitating the proof of Theorem 2 one arrives at a better upper bound (cf. Exercise 8):

$$R_k^{(r)}(s_1, \ldots, s_k) \le R_k^{(r-1)}(R_k^{(r)}(s_1 - 1, s_2, \ldots, s_k), \ldots, R_k^{(r)}(s_1, \ldots, s_{k-1}, s_k - 1)) + 1.$$

§2 Monochromatic Subgraphs

Let H_1 and H_2 be arbitrary graphs. Given n, is it true that every red-blue colouring of the edges of K^n contains a red H_1 or a blue H_2? Since H_i is a subgraph of K^{s_i}, where $s_i = |H_i|$, the answer is clearly "yes" if $n \geq R(s_1, s_2)$. Denote by $r(H_1, H_2)$ the smallest value of n that will ensure an affirmative answer. Note that this notation is similar to the one introduced earlier: $R(s_1, s_2) = r(K^{s_1}, K^{s_2})$. Clearly $r(H_1, H_2) - 1$ is the maximal value of n for which there is a graph G of order n such that $H_1 \not\subset G$ and $H_2 \not\subset \bar{G}$.

The numbers $r(H_1, H_2)$, sometimes called *generalized Ramsey numbers*, have been investigated fairly extensively in recent years. We shall determine $r(H_1, H_2)$ for some simple pairs (H_1, H_2).

Theorem 4. *Let T be a tree of order t. Then $r(K^s, T) = (s - 1)(t - 1) + 1$.*

Proof. The graph $(s - 1)K^{t-1}$ does not contain T, its complement, $K_{s-1}(t - 1)$, does not contain K^s, so $r(K^s, T) \geq (s - 1)(t - 1) + 1$.

Let now G be a graph of order $(s - 1)(t - 1) + 1$ whose complement does not contain K^s. Then $\chi(G) \geq \lceil n/(s - 1) \rceil = t$ so it contains a critical subgraph H of minimal degree at least $t - 1$ (see Theorem 1 of Chapter V). It is easily seen that H contains (a copy of) T. Indeed, we may assume that $T_1 \subset H$, where $T_1 = T - x$ and x is an endvertex of T, adjacent to a vertex y of T_1 (and of H). Since y has at least $t - 1$ neighbours in H, at least one of its neighbours, say z, does not belong to T_1. Then the subgraph of H spanned by T_1 and z clearly contains (a copy of) T.

As we know very little about $r(K^s, K^t)$, it is only to be expected that $r(G_1, G_2)$ has been calculated mostly in the cases when both G_1 and G_2 are sparse (have few edges compared to their orders), e.g., when $G_1 = sH_1$ and $G_2 = tH_2$. The following simple lemma shows that for fixed H_1 and H_2 the function $r(sH_1, tH_2)$ is at most $s|H_1| + t|H_2| + c$, where c depends only on H_1 and H_2, and not on s and t,

Lemma 5. $r(G, H_1 \cup H_2) \leq \max\{r(G, H_1) + |H_2|, r(G, H_2)\}$. *In particular,* $r(sH_1, H_2) \leq r(H_1, H_2) + (s - 1)|H_1|$.

Proof. Let n be greater than the right hand side and suppose there is a red-blue colouring of K^n without a red G. Then $n \geq r(G, H_2)$ implies that there is a blue H_2. Remove it. Since $n - |H_2| \geq r(G, H_1)$, the remainder contains a blue H_1. Hence K^n contains a blue $H_1 \cup H_2$. □

Theorem 6. *If $s \geq t \geq 1$ then*

$$r(sK^2, tK^2) = 2s + t - 1.$$

PROOF. The graph $G = K^{2s-1} \cup E^{t-1}$ does not contain s independent edges and $\bar{G} = E^{2s-1} + K^{t-1}$ does not contain t independent edges. Hence $r(sK^2, tK^2) \geq 2s + t - 1$.

Trivially $r(sK^2, K^2) = 2s$. We shall show that

$$r((s+1)K^2, (t+1)K^2) \leq r(sK^2, tK^2) + 3.$$

This is sufficient to complete the proof since then

$$\begin{aligned} r((s+1)K^2, (t+1)K^2) &\leq r(sK^2, tK^2) + 3 \leq r((s-1)K^2, (t-1)K^2) + 6 \\ &\leq r((s-t+1)K^2, K^2) + 3t = 2(s-t+1) + 3t \\ &= 2s + t + 2. \end{aligned}$$

Let G be a graph of order $n = r(sK^2, tK^2) + 3 \geq 2s + t + 1$. If $G = K^n$ then $G \supset (s+1)K^2$ and if $G = E^n$ then $\bar{G} \supset (t+1)K^2$. Otherwise there are three vertices, say x, y and z such that $xy \in G$, $xz \notin G$. Now either $G - \{x, y, z\}$ contains s independent edges and then xy can be added to them to form $s + 1$ independent edges of G or else $\bar{G} - \{x, y, z\}$ contains t independent edges and then xz can be added to them to form $t + 1$ independent edges of $\bar{G}$. □

Theorem 7. *If $s \geq t \geq 1$ and $s \geq 2$ then $r(sK^3, tK^3) = 3s + 2t$.*

PROOF. Let $G = K^{3s-1} \cup (K^1 + E^{2t-1})$. Then G does not contain s independent triangles and $\bar{G} = E^{3s-1} + (K^1 \cup K^{2t-1})$ does not contain t independent triangles. Hence $r(sK^3, tK^3)$ is at least as large as claimed.

It is easily shown that $r(2K^3, K^3) = 8$ and $r(2K^3, 2K^3) = 10$ (Exercise 12). Hence repeated applications of Lemma 5 give

$$r(sK^3, K^3) \leq 3s,$$

and to complete the proof it suffices to show that for $s \geq 1, t \geq 1$ we have

$$r((s+1)K^3, (t+1)K^3) \leq r(sK^3, tK^3) + 5.$$

To see this let $n = r(sK^3, tK^3) + 5$ and consider a red-blue colouring of K^n. Select a monochromatic (say red) triangle K_r in K^n. If $K^n - K_r$ contains a red sK^3 then we are home. Otherwise $K^n - K_r$ contains a blue triangle K_b (it even contains a blue tK^3). We may assume that at least five of the nine $K_r - K_b$ edges are red. At least two of these edges are incident with a vertex of K_b, and, together with an edge of K_r they form a red triangle K_r^* meeting K_b. Since $K^n - K_r^* - K_b$ has $r(sK^3, tK^3)$ vertices, it contains either a red sK^3 or a blue tK^3. These are disjoint from both K_r^* and K_b, so K^n contains either a red $(s+1)K^3$ or a blue $(t+1)K^3$. □

By elaborating the idea used in the proofs of the previous two theorems we can obtain good bounds on $r(sK^p, tK^q)$, provided $\max(s, t)$ is large

compared to $\max(p, q)$. Let $p, q \geq 2$ be fixed and choose t_0 such that

$$t_0 \min\{p, q\} \geq 2r(K^p, K^q).$$

Put $C = r(t_0 K^p, t_0 K^q)$.

Theorem 8. *If* $s \geq t \geq 1$ *then*

$$ps + (q-1)t - 1 \leq r(sK^p, tK^q) \leq ps + (q-1)t + C.$$

PROOF. The graph $K^{ps-1} \cup E^{(q-1)t-1}$ shows the first inequality. As in the proofs of the previous theorems, we fix $s - t$ and apply induction on t. By Lemma 5 we have

$$r(sK^p, tK^q) \leq (s-t)p + r(tK^p, tK^q) \leq ps + C,$$

provided $t \leq t_0$. Assume now that $t \geq t_0$, and the second inequality holds for s, t.

Let G be a graph of order $n = p(s+1) + (q-1)(t+1) + C + 1$ such that $G \not\supset (s+1)K^p$ and $\bar{G} \not\supset (t+1)K^q$. Suppose no K^p of G and K^q of $\bar{G}$ have a vertex in common. Denote by V_p the set of vertices of G that are in K^p subgraphs and put $V_q = V - V_p$, $n_p = |V_p|$, $n_q = |V_q|$. No vertex $x \in V_q$ is joined to more than $r(K^{p-1}, K^q)$ vertices of V_p since otherwise either there is a K^p of G containing x or else there is a K^q of $\bar{G}$ containing vertices of V_p. Similarly every vertex $y \in V_q$ is joined to all but at most $r(K^p, K^{q-1})$ vertices of V_p. Hence

$$n_q r(K^{p-1}, K^q) + n_p r(K^p, K^{q-1}) \geq n_p n_q.$$

However, this is impossible, since $n_p \geq sp$, $n_q \geq tq$ so $n_p > 2r(K^{p-1}, K^q)$ and $n_q > 2r(K^p, K^{q-1})$. Therefore we *can* find a K^p of G and a K^q of $\bar{G}$ with a vertex in common.

When we omit the $p + q - 1$ vertices of these subgraphs, we find that the remainder H is such that $H \not\supset sK^p$ and $\bar{H} \not\supset tK^q$. However, $|H| = ps + (q-1)t + C + 1$, so this is impossible. □

One is inclined to imagine that the various Ramsey theorems hold for finite graphs because the graph whose edges we colour is K^n and not some sparse graph with few edges.

Let us consider now another generalization of the original Ramsey problem for finite graphs. Let $H, G_1, G_2, \ldots, G_k$ be graphs. Denote by

$$H \rightarrow (G_1, G_2, \ldots, G_k)$$

the following statement: for every colouring of the edges of H with colours $c_1, c_2, \ldots, c_k$ there is an index i such that H contains a subgraph isomorphic to G_i whose edges are of colour c_i. If $G_i = G$ for each i then instead of $H \rightarrow (G_1, G_2, \ldots, G_k)$ we usually write $H \rightarrow (G)_k$. Note that

$$r(G_1, G_2) = \min\{n: K^{n+1} \rightarrow (G_1, G_2)\} = \max\{n: K^n \not\rightarrow (G_1, G_2)\}.$$

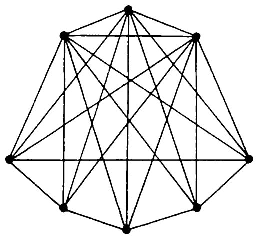


Figure VI.1. The graph $C^3 + C^5$.

Clearly it is harder to find a monochromatic subgraph in a sparse graph than in a complete graph and for given G_1 and G_2 it may not be easy to find a sparse graph H (sparse in some given sense) such that $H \to (G_1, G_2)$. Thus one may not find immediately a graph H without a K^6 for which $H \to (K^3, K^3)$. (Such a graph is shown in Figure VI.1). Going even further, one may ask the following difficult question: is there a graph H with $\mathrm{cl}(H) = 3$ (so without a K^4) such that $H \to (K^3, K^3)$? Note that Theorem 1 does *not* guarantee the existence of such an H. To conclude the section we only state a deep theorem of Nešetřil and Rödl showing that the answers to the questions of this type are in the affirmative.

Theorem 9. *Given graphs $G_1, G_2, \ldots, G_k$ there is a graph H such that* $\mathrm{cl}(H) = \max_{1 \le i \le k} \mathrm{cl}(G_i)$ *and* $H \to (G_1, G_2, \ldots, G_k)$. □

§3 Ramsey Theorems in Algebra and Geometry

Given an algebraic or geometric object and a collection $\mathscr{P}$ of finite subsets of the elements, is it true that whenever we colour the elements with k colours (i.e., partition them into k classes), at least one of the colour classes contains at least one member of $\mathscr{P}$? In this section we discuss some questions of this type for Euclidean spaces, finite vector spaces, various semigroups and objects belonging to rather general categories. We prove only one or two simple results, since the proofs of the deep results are well beyond the scope of these notes, though the statements themselves are very easily understood.

Let $\mathscr{P}$ be a collection of finite subsets of a set M. In accordance with the notation used in the previous section, $M \to (\mathscr{P})_k$ means that in every k-colouring of M there is a monochromatic set $P \in \mathscr{P}$, i.e., if $M = M_1 \cup \cdots \cup M_k$ then $P \subset M_i$ for some $i \in [1, k] = \{1, \ldots, k\}$ and $P \in \mathscr{P}$. When discussing whether $M \to (\mathscr{P})_k$ holds or not, the following compactness theorem enables us to replace M by a finite subset of it.

Theorem 10. $M \to (\mathscr{P})_k$ *iff there is a finite set* $X \subset M$ *such that* $X \to (\mathscr{P})_k$.

PROOF. If $M \not\to (\mathscr{P})_k$ then clearly $X \not\to (\mathscr{P})_k$ for every finite set X.

A k-colouring of M is a point of the space $[1, k]^M$; for $c \in [1, k]^M$ and $x \in M$ the projection of c into the x component, $\pi_x(c)$, is exactly $c(x) \in [1, k] = \{1, 2, \ldots, k\}$, the colour assigned by c to x. As almost always, we give $[1, k]$ the discrete topology and $[1, k]^M$ the product topology. By the familiar Tychonov theorem from general topology, the space $[1, k]^M$ is compact since it is a product of compact spaces.

Given a finite set $X \subset M$, let $N(X)$ be the set of colourings in which X does not contain a monochromatic set $P \in \mathscr{P}$. If $c \notin N(X)$ and $d \in [1, k]^M$ agrees with c on X (that is $d|X = c|X$) then $d \notin N(X)$. Therefore $[1, k]^M - N(X)$ is open and so $N(X)$ is closed. (In fact, $N(X)$ is trivially clopen.)

Let us turn now to the proof of the second implication. Assume that $X \not\to (\mathscr{P})_k$ for every finite set $X \subset M$. This means exactly that $N(X) \neq \emptyset$ for every finite set X. Since $N(X) \cap N(Y) \supset N(X \cup Y)$, the system $\{N(X): X \subset M \text{ is finite}\}$ of closed sets has the finite intersection property. The compactness of $[1, k]^M$ implies that $\cap_X N(X) \neq \emptyset$. Every $c \in \cap_X N(X)$ is a k-colouring of M without a monochromatic set $P \in \mathscr{P}$. □

Let L be a finite set of points in $\mathbb{R}^m$, the m-dimensional Euclidean space. As another variant of the notation used so far, $\mathbb{R}^n \to (L)_k$ means that in every k-colouring of $\mathbb{R}^n$ there is a monochromatic L' *congruent* to L. We say that L is *Ramsey* if for every k there is an n such that $\mathbb{R}^n \to (L)_k$.

Theorem 11. *Let P be a pair of points at distance* 1 *apart. Then*

$$\mathbb{R}^2 \to (P)_3 \quad \textit{but} \quad \mathbb{R}^2 \not\to (P)_7.$$

PROOF. Figure VI.2 shows the first assertion. For suppose that in a red-blue-yellow colouring of the seven points there is no monochromatic adjacent

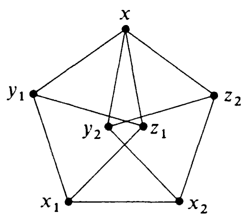


Figure VI.2. Adjacent points are at distance 1.

pair. We may assume that x is red. Then y_1, z_1 are blue and yellow so x_1 is red. Similarly x_2 is red, but x_1 and x_2 are adjacent.

Figure VI.3 shows that $\mathbb{R}^2 \to (P)_7$ is false. □

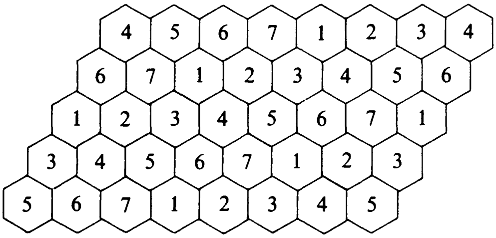


Figure VI.3. The hexagons have side a, $\frac{1}{2} < a < (4\sqrt{5} - 5)/10$.

It is interesting that we do not know the maximal value of k for which $\mathbb{R}^2 \to (P)_k$.

Theorem 12. *Let Q^2 be the (vertex set of a) unit square. Then $\mathbb{R}^6 \to (Q^2)_2$.*

PROOF. Consider a red-blue colouring of $\mathbb{R}^6$. Then, in particular, we have a red-blue colouring of the following fifteen points of $\mathbb{R}^6$: $x_{ij} = (x_{ij}^1, \ldots, x_{ij}^6)$, $1 \le i < j \le 6$, $x_{ij}^k = 0$ unless $k = i$ or j and $x_{ij}^i = x_{ij}^j = 1/\sqrt{2}$. Consider a K^6 with vertex set $\{v_1, \ldots, v_6\}$, and colour edge $v_i v_j (i < j)$ with the colour of x_{ij}. Since $r(C^4, C^4) = 6$ (the easy proof of this is left to the reader as Exercise 13), by symmetry we may assume that the edges $v_1 v_2$, $v_2 v_3$, $v_3 v_4$ are all red. It is easily checked that the points x_{12}, x_{23}, x_{34} and x_{14} form a unit square and this square is, of course, red. □

Given $L_1 \subset \mathbb{R}^{m_1}$ and $L_2 \subset \mathbb{R}^{m_2}$, we define $L_1 \times L_2 \subset \mathbb{R}^{m_1 + m_2}$ by $L_1 \times L_2 = \{(x_1, \ldots, x_{m_1+m_2}) : (x_1, \ldots, x_{m_1}) \in L_1, (x_{m_1+1}, \ldots, x_{m_1+m_2}) \in L_2\}$.

Theorem 13. *If L_1 and L_2 are Ramsey then so is $L_1 \times L_2$.*

PROOF. Given k, there is an m such that $\mathbb{R}^m \to (L_1)_k$. Hence by the compactness theorem there is a *finite* subset $X \subset \mathbb{R}^m$ such that $X \to (L_1)_k$. Put $l = k^{|X|}$. Since L_2 is Ramsey, there is an n such that $\mathbb{R}^n \to (L_2)_l$. Every k-colouring of $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$ restricts to a k-colouring $c: X \times \mathbb{R}^n \to [1, k] = \{1, 2, \ldots, k\}$. This gives an $l = k^{|X|}$-colouring of $\mathbb{R}^n$ with colours $[1, k]^X$: simply colour $y_0 \in \mathbb{R}^n$ with the function $f \in [1, k]^X$ that tells us how the points of $X \times \{y_0\}$ are coloured: $f(y_0) = c(x, y_0)$. In this l-colouring there is a monochromatic $L_2 \subset \mathbb{R}^n$, i.e., an $L_2 \subset \mathbb{R}^n$ such that for $x \in X$, and $y_1, y_2 \in L_2$ the points $(x, y_1), (x, y_2)$ have the same colour (in the original k-colouring). Assigning this common colour to x, we obtain a k-colouring of X. It has a monochromatic $L_1 \subset X$, providing us with a monochromatic $L_1 \times L_2$. □

A *brick* in $\mathbb{R}^n$ is a set congruent to

$$B = \{(\varepsilon_1 a_1, \varepsilon_2 a_2, \ldots, \varepsilon_n a_n) : \varepsilon_i = 0 \text{ or } 1\},$$

where $a_i > 0$. A unit simplex in $\mathbb{R}^k$ shows that $\mathbb{R}^k \to (\{0, 1\})_k$ so every 2-point set is Ramsey. Hence Theorem 13 has the following consequence.

Theorem 14. *Every subset of a brick is Ramsey.* □

Call a set $L \subset \mathbb{R}^m$ *spherical* if it is imbeddable in some sphere (of arbitrary dimension and radius). It is somewhat more complicated to show that if a set is not spherical then it is not Ramsey (Exercise 20). It is rather curious that though there are many sets (and many very simple sets, e.g., an obtuse triangle) that are spherical but not subsets of bricks, it is not known about any of them whether they are Ramsey or not.

The following very significant extension of Ramsey's theorem was proved by Hales and Jewett.

Theorem 15. *Let S be a commutative infinite semigroup. Let L be a finite subset of S and let* $\mathscr{P} = \{a + nL : a \in S, n \in \mathbb{N}\}$, *where* $nL = \{nx : x \in L\}$. *Then*

$$S \to (\mathscr{P})_k \quad \textit{for every } k.$$ □

The proof is too difficult to present in these notes. However, we state this theorem partly because it is the cornerstone of very general results proved by Graham, Leeb and Rothschild, asserting that certain categories are "Ramsey", and partly because we shall prove a special case of this theorem. The next theorem was conjectured by Rota; it is one of the most striking consequences of the results about Ramsey categories.

Theorem 16. *Given a finite field F and natural numbers n and k, there is a natural number N such that if* F^N, *the N-dimensional vector space over F, is coloured with k colours then* F^N *contains a monochromatic n-dimensional affine subspace.*

Furthermore, given a finite field F and natural numbers n, m and k, there is an M such that if the set of k-dimensional subspaces of F^M *is coloured with k colours, then there is an n-dimensional subspace in which all k-dimensional subspaces have the same colour.* □

The classical theorem of van der Waerden (Theorem 17 below) is a special case of Theorem 15. Here we present it with a proof due to Graham and Rothschild, since the proof gives back the flavour of the arguments used in the deepest results of the theory. As customary, we write $\mathbb{N}$ for the set of natural numbers and $[m, n]$ for $\{m, m + 1, \ldots, n\}$.

Theorem 17. *Given $p, k \in \mathbb{N}$, there exists an integer $W(p, k)$ such that in any k-colouring of $[1, W(p, k)]$ there is a monochromatic arithmetic progression consisting of p terms.*

PROOF. Let l and m be natural numbers. Two m-tuples $(x_1, \dots, x_m)$, $(y_1, \dots, y_m) \in [0, l]^{(m)}$ are said to be *l-equivalent* if they agree up to and including the last occurrences of l, with the convention that any two sequences not containing l are equivalent. We write $S(l, m)$ for the following statement:

Given $k \in \mathbb{N}$, there exists $W(l, m, k)$ such that for every function

$$c:[1, W(l, m, k)] \to [1, k]$$

there exist $a, d_1, \dots, d_m \in \mathbb{N}$ such that

$$a + l \sum_1^m d_i \le W(l, m, k) \quad \textit{and} \quad c\left(a + \sum_1^m x_i d_i\right)$$

is constant on each l-equivalence class of $[0, l]^{(m)}$.

Note that van der Waerden's theorem is exactly the assertion $S(p, 1)$, but the theorem holding for every p is equivalent to the truth of $S(l, m)$ for every l and m. We shall prove $S(l, m)$ by induction, starting with the trivial assertion $S(1, 1)$ and proving two different induction steps.

(i) *If $S(l, m)$ holds for some $m \ge 1$ then so does $S(l, m + 1)$.* To see this, fix k and put $W = W(l, m, k)$, $W' = W(l, 1, k^W)$ and $W_j = [(j - 1)W + 1, jW]$, $1 \le j \le W'$. Let $c:[1, WW'] \to [1, k]$ be any k-colouring. This induces a k^W-colouring c' of $[1, W]$ that colours $j \in [1, W]$ with the function $f \in [1, k]^{[1, W]}$ describing the colouring of W_j: $f(w) = c((j - 1)W + w)$, $1 \le w \le W$ (cf. the proof of Theorem 3). By the choice of W', there exist a' and d', $a' + ld' \le W'$ such that $c'(a') = c'(a' + d') = c'(a' + (l - 1)d')$. Let us apply $S(l, m)$ to the interval $[(a' - 1)W + 1, a'W]$, k-coloured with c. Then there exist $a, d_1, \dots, d_m \in \mathbb{N}$ such that

$$(a' - 1)W + 1 \le a \le a + l \sum_1^m d_i \le a'W$$

and $c(a + \sum_1^m x_i d_i)$ is constant on l-equivalence classes. Put $d_{m+1} = d'W$. Then $a, d_1, \dots, d_{m+1}$ are such that

$$a + l \sum_1^{m+1} d_i \le WW'$$

and

$$c\left(a + \sum_1^{m+1} x_i d_i\right)$$

is constant on each l-equivalence class of $[0, l]^{(m+)}$. (See Figure VI.4.)

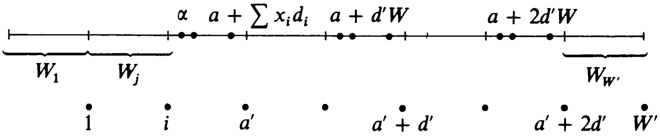


Figure VI.4. Illustration to part (i). The colour of $j \in [1, W']$ tells us how W_j is coloured.

(ii) *If $S(l, m)$ holds for every m then $S(l+1, 1)$ also holds.* This is an almost immediate consequence of the pigeonhole principle. For let $c: [W(l, k, k)] \to [1, k]$ be given. Then there exist $a, d_1, \ldots, d_k \in \mathbb{N}$ such that $a + l\sum_1^k d_i \le W(l, k, k)$ and $c(a + \sum_1^k x_i d_i)$ is constant on l-equivalence classes. The colours $c(a + \sum_1^s ld_i)$, $s = 0, 1, \ldots, k$ cannot be all different so there exist s, t, $0 \le s < t \le k$, such that

$$c\left(a + \sum_1^s ld_i\right) = c\left(a + \sum_1^t ld_i\right).$$

Then

$$c\left(\left(a + \sum_1^s ld_i\right) + j\sum_{s+1}^t d_i\right)$$

is constant for $j \in [0, j]$.

Since $S(1, 1)$ is trivial, by induction $S(l, m)$ is true for every l and m. □

Van der Waerden's theorem has many interesting and deep extensions. Among others, Rado determined when a system

$$\sum_{j=1}^n a_{ij}x_j = 0, \qquad i = 1, \ldots, m,$$

of linear equations with integer coefficients has a monochromatic solution in every colouring of $\mathbb{N}$ with finitely many colours. As a particular case of Rado's theorem and the compactness theorem one obtains the following result. Given integers k and n, there exists $N = N(k, n)$ such that if $[1, N]$ is k-coloured then there is a set A of n natural numbers such that $\sum_{a \in A} a \le N$ and all the sums $\sum_{b \in B} b$, $\emptyset \ne B \subset A$ have the same color. At the end of the next section we discuss a beautiful extension of this to infinite sets.

§4 Subsequences

Let (f_n) be a sequence of functions on a space T. Then we can find an infinite subsequence $(f_{n'})$ such that one of the following two alternatives holds:

a. if $(f_{n_i'})$ is any subsequence of $(f_{n'})$ then $\sup |\sum_1^N f_{n_i'}| \ge 1/N$ for every $N \ge 1$,
b. if $(f_{n_i'})$ is any subsequence of $(f_{n'})$ then $\sup |\sum_1^N f_{n_i'}| < 1/N$ for some $N \ge 1$.

This rather difficult assertion about sequences of functions is, in fact, an immediate consequence of a Ramsey type result about infinite sets.

As usual, 2^M denotes the set of subsets of M, $M(r)$ is the set of r-tuples of M and $M^{(\omega)}$ denotes the set of countably infinite subsets of M. $\mathbb{N}$ is the set of natural numbers and $[1, n] = \{1, 2, \ldots, n\}$. A family $\mathscr{F} \subset 2^{\mathbb{N}}$ is called *Ramsey* iff there exists an $M \in N^{(\omega)}$ such that either $M^{(\omega)} \subset \mathscr{F}$ or $M^{(\omega)} \subset 2^{\mathbb{N}} - \mathscr{F}$.

Of course, $2^{\mathbb{N}}$ can be identified with the cartesian product $\prod_{i \in \mathbb{N}} T_n$, where $T_n = \{0, 1\}$. We give T_n the discrete topology and the product $2^{\mathbb{N}}$ the product topology. In this topology $2^{\mathbb{N}}$ is a compact Hausdorff space. A weak form of a theorem due to Galvin and Prikry states that open subsets of $2^{\mathbb{N}}$ are Ramsey. (This is easily seen to imply the above assertion about sequences of functions.) To prove this result it is convenient to use the notation and terminology introduced by Galvin and Prikry. M, N, A and B are infinite subsets of $\mathbb{N}$. X, Y are finite subsets of N. $X < a$ means that $x < a$ for every $x \in X$; $X < M$ means that $X < m$ for every $m \in M$. An *M-extension* of X is a set of the form $X \cup N$ where $X < N$ and $N \subset M$. Let us fix now a family $\mathscr{F} \subset 2^{\omega}$. We say that *$M$ accepts X* if every M-extension of X belongs to $\mathscr{F}$; *M rejects X* if no $N \subset M$ accepts X.

Lemma 18. *If $\mathbb{N}$ rejects $\varnothing$ then there exists an $M \in N^{(\omega)}$ that rejects every $X \subset M$.*

Proof. Note first that there is an M_0 such that every $X \subset M_0$ is either accepted or rejected by M_0. Indeed, put $N_0 = \mathbb{N}$, $a_0 = 1$. Suppose that we have defined $N_0 \supset N_1 \supset \cdots \supset N_k$ and $a_i \in N_i - N_{i+1} 0 \le i \le k - 1$. Pick $a_k \in N_k$. If $N_k - \{a_k\}$ rejects $\{a_0, \ldots, a_k\}$ then put $N_{k+1} = N_k - \{a_k\}$, otherwise let N_{k+1} be an infinite subset of $N_k - \{a_k\}$ that accepts $\{a_0, \ldots, a_k\}$. Then $M_0 = \{a_0, a_1, \ldots\}$ will do.

By assumption M_0 rejects $\varnothing$. Suppose now that we have chosen $b_0, b_1, \ldots, b_{k-1}$ such that M_0 rejects every $X \subset \{b_0, b_1, \ldots, b_{k-1}\}$. Then M_0 can not accept infinitely many sets of the form $X \cup \{c_j\}, j = 1, 2, \ldots,$ since otherwise $\{c_1, c_2, \ldots\}$ accepts X. Hence M_0 rejects all but finitely many sets of the form $X \cup \{c\}$. As there are only 2^k choices for X, there exists a b_k such that M_0 rejects every $X \subset \{b_0, b_1, \ldots, b_k\}$. By construction the set $M = \{b_0, b_1, \ldots\}$ has required property. □

Theorem 19. *Every open subset of $2^{\mathbb{N}}$ is Ramsey.*

Proof. Let $\mathscr{F} \subset 2^{\mathbb{N}}$ be open and assume that $A^{(\omega)} \not\subset \mathscr{F}$ for every $A \in N^{(\omega)}$, i.e., $\mathbb{N}$ rejects $\varnothing$. Let M be the set whose existence is guaranteed by Lemma 18. If $M^{(\omega)} \not\subset 2^{\omega} - \mathscr{F}$, let $A \in M^{(\omega)} \cap \mathscr{F}$. Since $\mathscr{F}$ is open, it contains a neighbourhood of A so there is an integer $a \in A$ such that if $B \cap \{1, 2, \ldots, a\} = A \cap \{1, 2, \ldots, a\}$ then $B \in \mathscr{F}$. But this implies that M accepts $A \cap \{1, 2, \ldots, a\}$, contrary to the choice of M. Hence $M^{(\omega)} \subset 2^{\omega} - \mathscr{F}$, proving that $\mathscr{F}$ is Ramsey. □

Denote by $X^{(<\omega)}$ the family of *finite* subsets of X. A family $\mathscr{F}_0 \subset \mathbb{N}^{(<\omega)}$ is *dense* if $\mathscr{F}_0 \cap M^{(<\omega)} \neq \varnothing$ for every $M \in N^{(\omega)}$ and it is *thin* if no member of $\mathscr{F}_0$ is an initial segment of another member (that is if $X < Y$ implies $X \notin \mathscr{F}_0$ or $X \cup Y \notin \mathscr{F}_0$).

Corollary 20. *Let $\mathscr{F}_0 \subset N^{(<\omega)}$ be dense. Then there is an $M \in N^{(\omega)}$ such that every $A \subset M$ has an initial segment belonging to $\mathscr{F}_0$.*

PROOF. Let $\mathscr{F} = \{F \subset \mathbb{N}: F$ has an initial segment belonging to $\mathscr{F}_0\}$. Then $\mathscr{F}$ is open so there is an $M \in \mathbb{N}^{(\omega)}$ such that either $M^{(\omega)} \subset \mathscr{F}$, in which case we are home, or else $M^{(\omega)} \subset 2^{\mathbb{N}} - \mathscr{F}$. The second alternative cannot hold since it would imply $M^{(<\omega)} \cap \mathscr{F}_0 = \varnothing$. □

This corollary enables us to deduce an extension of the original Ramsey theorem for infinite sets (Theorem 3).

Corollary 21. *Let $\mathscr{F}_0 \subset N^{(<\omega)}$ be a thin family. Then for any k-colouring of $\mathscr{F}_0$ there is an infinite set $A \subset \mathbb{N}$ such that all members of $\mathscr{F}_0$ contained in A have the same colour.*

PROOF. It suffices to prove the result for $k = 2$. Consider a red and blue colouring of $\mathscr{F}_0$: $\mathscr{F}_0 = \mathscr{F}_{\text{red}} \cup \mathscr{F}_{\text{blue}}$. If $\mathscr{F}_{\text{red}}$ is dense then let M be the set guaranteed by Corollary 20. For every $F \in \mathscr{F}_0 \cap 2^M$ there is an infinite set $N \subset M$ with initial segment F. Since $\mathscr{F}_0$ is thin, F is the unique initial segment of N that belongs to $\mathscr{F}_0$. Hence $F \in \mathscr{F}_{\text{red}}$, so every member of $\mathscr{F}_0$ contained in M is red.

On the other hand if $\mathscr{F}_{\text{red}}$ is not dense then $2^M \cap \mathscr{F}_{\text{red}} = \varnothing$ for some infinite set M. Hence $2^M \cap \mathscr{F}_0 \subset \mathscr{F}_{\text{blue}}$. □

Let us now turn to the result concerning monochromatic sums we promised at the end of the previous section. This beautiful result, conjectured by Graham and Rothschild and first proved by Hindman, is not very near to the results given in this section, but the striking proof given by Glazer illustrates the rich methods that can be applied in infinite Ramsey theory.

Theorem 22. *For any k-colouring of $\mathbb{N}$ there is an infinite set $A \subset \mathbb{N}$ such that all sums $\sum_{x \in X} x$, $\varnothing \neq X \subset A$, have the same colour.* □

We shall not give a detailed proof but only sketch one for those who are (at least rather vaguely) familiar with ultrafilters on $\mathbb{N}$ and know that the set $\beta\mathbb{N}$ of all ultrafilters is a compact topological space (the Stone–Čech compactification of $\mathbb{N}$ with, of course, the discrete topology). The proof which is due to Glazer, is at least as beautiful as the theorem and it is considerably more surprising.

Let us recall that a filter $\mathscr{F}$ on $\mathbb{N}$ is a collection of subsets of $\mathbb{N}$ such that (i) if $A, B \in \mathscr{F}$, then $A \cap B \in \mathscr{F}$, (ii) if $A \in \mathscr{F}$ and $A \subset B$ then $B \in \mathscr{F}$ and (iii) $\mathscr{F} \neq 2^{\mathbb{N}}$, that is $\varnothing \notin \mathscr{F}$. Zorn's lemma implies that every filter is contained in a maximal filter, called *ultrafilter*. If $\mathscr{U}$ is an ultrafilter then for every $A \subset \mathbb{N}$ either $A \in \mathscr{U}$ or else $\mathbb{N} - A \in \mathscr{U}$. This implies that every ultrafilter $\mathscr{U}$ defines a finitely additive 0-1 measure m on $2^{\mathbb{N}}$:

$$m(A) = \begin{cases} 1 & \text{if } A \in \mathscr{U}, \\ 0 & \text{if } \mathbb{N} - A \in \mathscr{U}. \end{cases}$$

Conversely, clearly every finitely additive 0-1 measure on $2^{\mathbb{N}}$ defines an ultrafilter. If there is a *finite* set of measure 1 then one of the elements, say a, of that set has also measure 1, and so $\mathscr{U} = \{A \subset \mathbb{N} : a \in A\}$. These ultrafilters are called *principal*. Not every ultrafilter is principal: the ultrafilters containing the filter $\mathscr{F} = \{A \subset N : \mathbb{N} - A \text{ is finite}\}$ are *not* principal.

That ultrafilters can be useful in proofs of Ramsey theorems can be seen from the following very simple proof of the case $r = 2$ of Theorem 3. Let $\mathscr{U}$ be a non-principal ultrafilter. Let $N^{(2)} = P_1 \cup P_2 \cup \cdots \cup P_k$. For $n \in \mathbb{N}$ let $A_i^{(n)} = \{m : (n, m) \in P_i\}$. Then exactly one of the sets $A_1^{(n)}, A_2^{(n)}, \ldots, A_k^{(n)}$ belongs to $\mathscr{U}$, say the set $A_{c(n)}^{(n)}$. Now with $B_i = \{n : c(n) = i\}$ we have $\mathbb{N} = B_1 \cup \cdots \cup B_k$ so again exactly one of these sets, say B_j, belongs to $\mathscr{U}$. Finally, pick $a_1 \in B_j$, $a_2 \in B_j \cap A_j^{(a_1)}$, $a_3 \in B_j \cap A_j^{(a_1)} \cap A_j^{(a_2)}$, etc. With $A = \{a_1, a_2, \ldots\}$ we have $A^{(2)} \subset P_j$.

Let us turn at last to the sketch of Glazer's proof of Theorem 22. Let us define an addition on $\beta\mathbb{N}$ by

$$\mathscr{U} + \mathscr{V} = \{A \subset \mathbb{N} : \{n \in \mathbb{N} : A - n \in \mathscr{U}\} \in \mathscr{V}\}$$

where $\mathscr{U}, \mathscr{V} \in \beta\mathbb{N}$ and $A - n = \{a - n : a \in A, a > n\}$.

With some effort one can check that $\mathscr{U} + \mathscr{V}$ is indeed an ultrafilter and that with this addition $\beta\mathbb{N}$ becomes a semigroup. Furthermore, the semigroup operation is right-continuous, i.e. for a fixed $\mathscr{P} \in \beta\mathbb{N}$ the map $\beta\mathbb{N} \to \beta\mathbb{N}$, given by $\mathscr{U} \to \mathscr{P} + \mathscr{U}$, is continuous. By applying a short and standard topological argument we see that the properties above imply that $\beta\mathbb{N}$ has an idempotent element, that is an element $\mathscr{P}$ with $\mathscr{P} + \mathscr{P} = \mathscr{P}$. This $\mathscr{P}$ is nonprincipal since if $\{p\} \in \mathscr{P}$ then $\{2p\} \in \mathscr{P} + \mathscr{P}$ so $\{p\} \notin \mathscr{P} + \mathscr{P}$.

Let now $A \in \mathscr{P}$. Then by the definition of the addition

$$A^* = \{n \in N : A - n \in P\} \in \mathscr{P}.$$

Thus if $a \in A \cap A^*$ then $B = (A - a) \cap (A - \{a\}) \in \mathscr{P}$. (We could replace A by $A - \{a\}$ since $\mathscr{P}$ is not principal.) Hence for every $A \in \mathscr{P}$ there exist $a \in A$ and $B \subset A - \{a\}$ such that $B \in \mathscr{P}$ and $a + B \subset A$.

Of course, this ultrafilter P has nothing to do with any colouring of N. However, just as any non-principal ultrafilter enabled us to find a monochromatic infinite set in a direct way, this idempotent $\mathscr{P}$ enables us to find an appropriate infinite set. Let $N = C_1 \cup \cdots \cup C_k$ be the decomposition

of N into colour classes. Exactly one of these colour classes, say C_i, belongs to $\mathscr{P}$. Put $A_1 = C_i$. Select $a_1 \in A_1$ and $A_2 \in \mathscr{P}$, $A_2 \subset A_1 - \{a_1\}$ such that $a_1 + A_2 \subset A_1$. Then select $a_2 \in A_2$ and $A_3 \in \mathscr{P}$, $A_3 \subset A_2 - \{a_2\}$ such that $a_2 + A_3 \subset A_2$, etc. The set $A = \{a_1, a_2, \dots\}$ clearly has the required property: every infinite sum $\sum_{x \in X} x$, $X \subset A$, has colour i. □

Finally, it should be emphasised that the infinite Ramsey results presented in this section form only the tip of an iceberg: the Ramsey theory of infinite sets, called *partition calculus*, is an essential and very cultivated branch of set theory, and it has a huge literature.

The foundations of partition calculus were laid down over twenty years ago by Erdös and Rado, who also introduced the arrow notation, used throughout the chapter, to formulate assertions about large cardinals.

Exercises

1. Show the equivalence of Theorem 3 and $R_k^{(r)}(s_1, \dots, s_k) < \infty$,
   (i) directly,
   (ii) using Theorem 10.
2. Prove that $R(3, 4) = 9$.

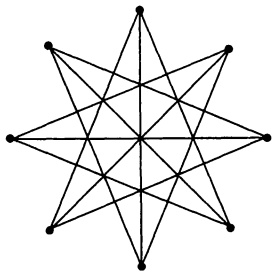


Figure VI.5. A graph showing $R(3, 4) \geq 9$.

3. Extending the construction in Figure VI.5, find for each $t \geq 2$ a t-regular graph which shows that $R(3, t + 1) \geq 3t$.
4. By considering the graph with vertex set $\mathbb{Z}_{17}$ (the integers modulo 17) in which i is joined to j iff $i - j = \pm 1, \pm 2, \pm 4$ or ± 8, show that $R(4, 4) = 18$.
5. Prove that $R(3, 5) = 14$.
6. Let e be an edge of K^4. Show that $r(K^4 - e, K^4) = 11$ (cf. Exercise 4).
7. By considering the 3-colouring of K^{16} with vertex set $GF(16)$, the field of order 16, in which the colour of an edge ij depends on the coset of the cubic residues to which $i - j$ belongs, show that $R_3(3, 3, 3) = 17$. (You have to check that the graph is well defined.)
8. Establish the upper bound for $R_k^{(r)}(s_1, \dots, s_k)$ given at the end of §1.

9.⁻ Given $2 \le k \le n$, denote by $c_k(n)$ the maximal integer which is such that in every k-colouring of the edges of K^n we can find a connected monochromatic subgraph of order $c_k(n)$. Show that $c_2(n) = n$. (See Exercise 13 of Chapter I.)

10.⁻ Prove that $c_{n-1}(n) = 2$ if $n \ge 2$ is even and $c_{n-1}(n) = 3$ if $n \ge 3$ is odd. [*Hint.* Use Theorem 9 of Chapter I.]

11.⁺ Prove that

$$c_3(n) = \begin{cases} \dfrac{n}{2} + 1 & \text{if } n \equiv 2 \pmod 4, \\ \left\lfloor \dfrac{n}{2} \right\rfloor & \text{otherwise} \end{cases}$$

12. Check that $r(2K^3, K^3) = 8$ and $r(2K^3, 2K^3) = 10$.

13. Show that $r(C^4, C^4) = 6$.

14. By using the proof of Theorem 9 of Chapter IV prove that $r_k(C^4, C^4, \ldots, C^4) \le k^2 + k + 2$.

15.⁺ Using two-dimensional vector spaces over finite fields (cf. Theorem 10 of Chapter IV), show that

$$r_k(C^4, C^4, \ldots, C^4) = k^2 + O(k).$$

16. By considering $H_1 = P^5$ and $H_2 = K^{1,3}$, show that

$$r(H_1, H_2) \ge \min_{i=1,2} r(H_i, H_i)$$

need not hold.

17.⁺ Let H_p, H_q be graphs of order p and q, respectively and let $\mathrm{cl}(\overline{H}_p) = i$, $\mathrm{cl}(\overline{H}_q) = j$. Then there is a constant C depending only on p and q such that

$$ps + qt - \min\{si, tj\} - 2 \le r(sH_p, tH_q) \le ps + qt - \min\{si, tj\} + C.$$

[*Hint.* Find a red $K^{j(p-i)}$, say R, a blue $K^{i(q-j)}$, say B, and a set N of ij other vertices such that the RN edges are red and the BN edges are blue. Cf. the proof of Theorem 8.]

18. Prove that a set $\{x_0, x_1, \ldots, x_l\} \subset \mathbb{R}^n$ is non-spherical iff for some c_i

$$\sum_1^l c_i(x_i - x_0) = 0 \quad \text{and} \quad \sum_1^l c_i(|x_i|^2 - |x_0|^2) = b \ne 0.$$

19. Let $b, c_1, \ldots, c_l \in \mathbb{R}, b \ne 0$. Show that there exists an integer k and some k-colouring of $\mathbb{R}^1$ such that the equation

$$\sum_1^l c_i(x_i - x_0) = b$$

has no solution with $x_0, x_1, \ldots, x_l$ all the same colour.

20. Prove that every Ramsey set is spherical. [*Hint.* Given a non-spherical set $\{x_0, x_1, \ldots, x_l\} \subset \mathbb{R}^n$, find (c_i) and b as in Exercise 18. Choose a k and a k-colouring of $\mathbb{R}^1$ as in Exercise 19. Use this k-colouring to define a k-colouring of $\mathbb{R}^n$ by colouring $x \in \mathbb{R}^n$ with the colour of $|x|^2$.]

21. Let $f(n)$ be the minimal integer N such that whenever X is a set of N points in a plane, no three of which are collinear, X contains n points forming a convex n-gon. Show that $f(n) \leq R^{(4)}(5, n)$ for every $n \geq 4$. Can you give a better bound for $f(5)$?

22. Let S be an infinite set of points in the plane. Show that there is an infinite set A such that either A is contained in a line or no three points of A are collinear.

23. Show that $R_k(3, 3, \ldots, 3) \leq \lfloor ek! \rfloor + 1$.

24. Deduce from Exercise 23 the following theorem of Schur: If we partition the natural numbers not exceeding $ek!$ into k classes then the equation $x + y = z$ is solvable in at least one of the classes.

25. Show that there is an infinite set of natural numbers such that the sum of any two elements has an even number of prime factors.

26.$^+$ Show that there is a sequence $n_1 < n_2 < \cdots$ of natural numbers such that if $r \leq i_1 < i_2 < \cdots < i_r$ then $\sum_{j=1}^{r} n_{i_j}$ has an even number of prime factors iff r has an odd number of prime factors.

27. Define a graph with vertex set $[1, N]^{(2)}$ by joining $a < b$ to $b < c$. Show that this graph does not contain a triangle and its chromatic number tends to infinity with N. (See Exercise 11 of Chapter V.)

28. Let $g_1(x), g_2(x), \ldots, g_n(x)$ be bounded real functions and let $f(x)$ be another real function. Let ε and δ be positive constants. Suppose $\max_i (g_i(x) - g_i(y)) > \delta$ whenever $f(x) - f(y) > \varepsilon$. Prove that f is bounded.

Notes

The fundamental Ramsey theorems are in F. P. Ramsey, On a problem of formal logic, *Proc. Lond. Math. Soc.* (2) **30** (1930) 264–286. Theorems 4,7 and 8 are due to V. Chvátal, P. Erdös and J. Spencer and can be found in S. A. Burr, Generalized Ramsey theory for graphs—a survey, in *Graphs and Combinatorics* (R. Bari and F. Harary, eds.) Springer-Verlag, 1974, pp. 52–75. For some of the deep results of Nešetřil and Rödl, one of which is Theorem 9, see J. Nešetřil and V. Rödl, The Ramsey property for graphs with forbidden subgraphs, *J. Combinatorial Theory* Ser. B, **20** (1976) 243–249, and Partitions of finite relational and set systems, *J. Combinatorial Theory* Ser. A **22** (1977) 289–312.

The geometric Ramsey results of §3 can be found in P. Erdös, R. L. Graham, B. L. Rothschild, J. Spencer and E. G. Straus, Euclidean Ramsey theorems, I, *J. Combinatorial Theory* Ser. A, **14** (1973) 344–363. Theorem 15 is in A. Hales and R. I. Jewett, Regularity and positional games, *Trans. Amer. Math. Soc.* **106** (1963) 222–229. Rota's conjecture is proved in R. L. Graham and B. L. Rothschild, Ramsey's theorem for n-parameter sets, *Trans. Amer. Math. Soc.* **159** (1971) 257–292. The results concerning Ramsey

categories are given in R. L. Graham, K. Leeb and B. L. Rothschild, Ramsey's theorem for a class of categories, *Adv. in Math.* **8** (1972) 417–433 (Errata, ibid **10** (1973) 326–327). Van der Waerden's theorem is in B. L. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Archief voor Wiskunde* **15** (1927) 212–216. The simple proof we give is from R. L. Graham and B. L. Rothschild, A short proof of van der Waerden's theorem on arithmetic progressions, *Proc. Amer. Math. Soc.* **42** (1974) 385–386.

Rado's theorem, mentioned at the end of §3 is from R. Rado, Studien zur Kombinatorik, *Math. Zeitschrift* **36** (1933) 424–480.

The first results of §4 are in F. Galvin and K. Prikry, Borel sets and Ramsey's theorem, *J. Symbolic Logic* **38** (1973) 193–198; Glazer's proof of Hindman' theorem is in W. W. Comfort, Ultrafilters: some old and some new results, *Bull. Amer. Math. Soc.* **83** (1977) 417–455.

Exercise 5 is in R. E. Greenwood and A. M. Gleason, Combinatorial relations and chromatic graphs, *Canad. J. Math.* **7** (1955) 1–7, Exercise 9 in a remark of F. Galvin and Schur's theorem (Exercise 24) is in I. Schur, Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$, *Jahr. Deutsch. Math. Ver.* **25** (1916) 114–116.

CHAPTER VII
Random Graphs

"Give a 'good' lower bound on the Ramsey number $R(s, s)$, that is show that there exists a graph of large order such that neither the graph nor its complement contains a K^s". "Show that for every natural number k there is a k-chromatic graph which does not contain a triangle". It does not take long to realize that the constructions that seem to be demanded by questions like these are not easily come by. Later we show that for every k there is a graph with the latter property, but even for $k = 4$ our graph has at least 2^{32} vertices. This book does not contain a picture of such a graph. Indeed, the aim of this chapter is to show that in order to solve these problems we can use probabilistic methods to demonstrate the existence of the graphs without actually constructing them. (It should be noted that we never use more than the convenient *language* of probability theory, since all the probabilistic arguments we need can be replaced by *counting* the number of objects in various sets.) This phenomenon is not confined to graph theory and combinatorics; in the last decade or two probabilistic methods have been used with striking success in Fourier analysis, in the theory of function spaces, in number theory, in the geometry of Banach spaces, etc. However, there is no area where probabilistic (or counting) methods are more natural than in combinatorics.

In most cases we use one of two closely related probabilistic models. We start with a *fixed set of n distinguished* (labelled) *vertices*. Then either we choose every edge with some probability $p, 0 < p < 1$, independently of the choices of other edges or else we take all graphs of order n and size M, and consider them as points of a probability space, having equal probability. In the first case we shall write $\mathcal{G}(n, P(\text{edge}) = p)$ for the probability space, in the second case we write $\mathcal{G}(n, M)$. Of course in a more complicated model

we would choose the edges with given probabilities from several disjoint subsets of edges.

In the first two sections we shall show that a random graph can be a valuable weapon in attacking straightforward questions concerning graphs, including the ones mentioned above. However, random graphs are clearly of interest in themselves. A random graph can be viewed as an organism that develops by acquiring more and more edges as p or M increases. It is fascinating and rather surprising that a given property may appear rather suddenly at a certain stage of the development. In the last two sections we present some examples of this phenomenon.

§1 Complete Subgraphs and Ramsey Numbers—the Use of the Expectation

Let n, N and M be natural numbers, $0 \leq M \leq N = \binom{n}{2}$. We shall consider the set $\Omega = \mathcal{G}(n, M)$ of all graphs of size M with vertex set $V = \{1, 2, \dots, n\}$. Clearly Ω has $\binom{N}{M}$ elements; in particular, if $M = 0$ then $\Omega = \{E^n\}$ and if $M = N = \binom{n}{2}$ then $\Omega = \{K^n\}$. (We use M for the number of edges in order to emphasise that M tends to be of larger order than n. We cannot use e since the base of the natural logarithm occurs in many of our formulae.) For the sake of convenience we view Ω as a probability space, in which the points (i.e., graphs) have equal probability, $1/|\Omega|$. Then all graphic invariants occur as random variables on Ω, so we may talk of their expected value, standard deviation, etc. In our calculations we shall often make use of various estimates of binomial coefficients. For convenience we list them now. First recall Stirling's formula

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\, e^{\alpha/(12n)}, \tag{1}$$

where α depends on n but is between 0 and 1. The following estimates, in which $0 \leq x \leq x + y \leq a$ and $x \leq b \leq a$, are obtained by expanding the binomial coefficient and occasionally applying Stirling's formula (1).

$$\frac{1}{2\sqrt{a}}\left(\frac{a}{b}\right)^b \leq \frac{1}{2\sqrt{a}}\left(\frac{a}{a-b}\right)^{a-b}\left(\frac{a}{b}\right)^b \leq \binom{a}{b} \leq \left(\frac{a}{a-b}\right)^{a-b}\left(\frac{a}{b}\right)^b \leq \left(\frac{ea}{b}\right)^b, \tag{2}$$

$$\left(\frac{a-b-y}{a-x-y}\right)^y\left(\frac{b-x}{a-x}\right)^x \leq \binom{a-x-y}{b-x}\binom{a}{b}^{-1}$$
$$\leq \left(\frac{a-b}{a}\right)^y\left(\frac{b}{a}\right)^x \leq e^{-(b/a)y-(1-(b/a))x}. \tag{3}$$

As our first illustration of the use of the language of probability in graph theory we give a lower bound for the Ramsey number $R(s, t)$. This lower bound is based on the expected number of complete subgraphs of a given order. For $G \in \Omega$ denote by $X_s = X_s(G)$ the number of K^s subgraphs in G. Then X_s is an integer valued random variable on the probability space Ω.

Theorem 1. *The expected number of K^s subgraphs contained in a graph $G \in \Omega = \mathscr{G}(n, m)$ is*

$$E(X_s) = \binom{n}{s}\binom{N - \binom{s}{2}}{M - \binom{s}{2}}\binom{N}{M}^{-1}.$$

PROOF. As we noted earlier, $|\Omega| = \binom{N}{M}$. In order to calculate the expected number of K^s subgraphs in $G \in \Omega$, first we compute the number of graphs $G \in \Omega$ that contain a fixed complete subgraph K_0 of order s. If $K_0 \subset G$ then $\binom{s}{2}$ of the edges of G are determined (i.e., are in K_0) and the remaining $M - \binom{s}{2}$ edges have to be chosen from a set of $N - \binom{s}{2} = \binom{n}{2} - \binom{s}{2}$ edges. Thus

$$\binom{N - \binom{s}{2}}{M - \binom{s}{2}}$$

of the graphs $G \in \Omega$ contain K_0. As there are $\binom{n}{s}$ choices for K_0, the expected number of K^s subgraphs contained in $G \in \Omega$ is as claimed. □

Theorem 2. *If $s, t \geq 3$ then*

$$R(s, t) > \exp\left\{\frac{(s-1)(t-1)}{2(s+t)}\right\}.$$

In particular

$$R(s, s) > e^{(1/4)(s-1)^2/s}.$$

PROOF. Note first that the graph $G = (t-1)K^{s-1}$ does not contain a K^s and its complement, $K_{s-1}(t-1)$, does not contain a K^t, so $R(s, t) \geq (s-1)(t-1) + 1$. Therefore we may assume that

$$n = \left\lceil \exp\left\{\frac{(s-1)(t-1)}{2(s+t)}\right\}\right\rceil \geq (s-1)(t-1) + 1,$$

so, in particular $8 \leq s \leq t$.

Let

$$M = \left\lfloor \frac{s}{s+t} N \right\rfloor, \qquad M' = N - M < \frac{t}{s+t} N + 1,$$

let E_s be the expected number of K^s subgraphs in a graph of order n and size M, and, similarly, let E_t be the expected number of K^t subgraphs in the

complement of such graphs (i.e., in the graphs $G(n, M')$). Let us estimate the expression for E_s, given in Theorem 1, by using inequalities (2) and (3). We obtain

$$\log E_s \le s(\log n + 1 - \log s) - \frac{t}{s+t}\frac{s(s-1)}{2}$$
$$\le -\frac{s(s-1)}{2(s+t)} + 1 - s(\log s - 1) < -3.$$

Similarly

$$\log E_t \le t(\log n + 1 - \log t) - \frac{s}{s+t}\frac{t(t-1)}{2} + \frac{t(t-1)}{n(n-1)}$$
$$= -\frac{t(t-1)}{2(s+t)} + 1 - t(\log t - 1) + \frac{t(t-1)}{n(n-1)} < -3.$$

Hence $E_s + E_t < 1$ so there is a graph $G \in \Omega$ which does not contain a K^s and whose complement does not contain a K^t. Hence $R(s, t) \ge n$, as claimed. □

Clearly the bound given in Theorem 2 is very bad for $s = 3$ and large t. In §2 we shall sketch a more refined version of the proof above, that gives a much better bound for $R(3, t)$. However, first we apply a straightforward argument to obtain lower bounds in the problem of Zarankiewicz (cf. Theorems 8 and 10 in Chapter IV).

Theorem 3. *Let* $2 \le s \le n_1$, $2 \le t \le n_2$, $\alpha = (s-1)/(st-1)$ *and* $\beta = (t-1)/(st-1)$. *Then there is a bipartite graph* $G_2(n_1, n_2)$ *of size*

$$\left\lfloor \left(1 - \frac{1}{s!t!}\right) n_1^{1-\alpha} n_2^{1-\beta} \right\rfloor$$

that does not contain a $K(s, t)$ (*with* s *vertices in the first class and* t *vertices in the second class*).

PROOF. Let

$$\begin{aligned} n &= n_1 + n_2, \\ V_1 &= \{1, 2, \ldots, n_1\}, \\ V_2 &= \{n_1 + 1, n_1 + 2, \ldots, n_1 + n_2\}, \\ E &= \{ij\colon i \in V_1, j \in V_2\}, \\ M &= \lfloor n_1^{1-\alpha} n_2^{1-\beta} \rfloor. \end{aligned}$$

We shall consider the probability space $\Omega = \mathcal{G}(n, M; E)$ consisting of the $\binom{|E|}{M}$ graphs with vertex set $V = V_1 \cup V_2$ having exactly M edges from E and none outside E. (Note that this is *not* the probability space considered

in the previous theorems.) The expected number of $K(s, t)$ subgraphs contained in a graph $G \in \Omega$ is

$$E_{s,t} = \binom{n_1}{s}\binom{n_2}{t}\binom{|E| - st}{M - st}\binom{|E|}{M}^{-1},$$

since the first factor is the number of ways the first class of $K(s, t)$ can be chosen, the second factor is the number of ways the second class can be chosen and since $K(s, t)$ has st edges. By (3) we have

$$E_{s,t} \le \frac{1}{s!t!} n_1^s n_2^t \left(\frac{M}{n_1 n_2}\right)^{st} < \frac{1}{s!t!} n_1^{1-\alpha} n_2^{1-\beta}.$$

Thus there is a graph $G_0 \in \Omega$ that contains fewer than $(1/s!t!)n_1^{1-\alpha}n_2^{1-\beta}$ complete bipartite graphs $K(s, t)$. Omit one edge from each $K(s, t)$ in G_0. The obtained graph $G = G_2(n_1, n_2)$ has at least

$$\lfloor n_1^{1-\alpha} n_2^{1-\beta} \rfloor - \frac{1}{s!t!} n_1^{1-\alpha} n_2^{1-\beta} \ge \left\lfloor \left(1 - \frac{1}{s!t!}\right) n_1^{1-\alpha} n_2^{1-\beta} \right\rfloor$$

edges and contains no $K(s, t)$. □

By similar methods one can construct a graph of order n and size

$$\left\lfloor \frac{1}{2}\left(1 - \frac{1}{s!t!}\right) n^{2-(s+t-2)/(st-1)} \right\rfloor$$

that does not contain a $K(s, t)$ (see Exercise 4).

It is a fact of life that the answers produced with the aid of random graphs usually differ from the best possible ones by a factor of 2. For example, random graphs give $R(s, s) > 2^{s/2}$, whereas one would expect the actual value of $R(s, s)$ to be around 2^s. Similarly it is thought that there are graphs of order n with $c_t n^{2-(1/t)}$ edges containing no $K(t, t)$, where c_t is a positive constant depending only on t; random graphs show that $c_t n^{2-(2/t)}$ edges are possible. However, in both these cases random graphs produce the best known lower bounds.

§2 Girth and Chromatic Number—Altering a Random Graph

When looking for graphs with a certain property, we sometimes turn a set of graphs into a probability space and hope that most graphs in that space have the required property. If this does not happen, all is not lost, for we can try to alter the graph slightly to make it have that property. We shall illustrate this point with a particularly simple example. Throughout the

section, we work with the probability space $\Omega = \mathscr{G}(n, M)$ of all graphs of size M with a fixed set of n distinguished vertices.

Theorem 4. *Given natural numbers $\delta \geq 3$ and $g \geq 4$ there is a graph of order at most $(2\delta)^g$ whose minimum degree is at least δ and whose girth is at least g.*

PROOF. Let $\Omega = \mathscr{G}(n, M)$, where $n = (2\delta)^g$ and $M = \delta n$. Denote by $Z_l(G)$ the number of cycles of length l in a graph G; thus Z_l is a random variable on the space Ω. What is $E(Z_l)$, the expected value of Z_l? A cycle C_0 of length l determines l edges of a graph G containing it; the remaining $M - l$ edges of G have to be chosen from a set of $\binom{n}{2} - l = N - l$ edges. Since there are $\binom{n}{l}$ choices for the vertex set of C_0 and, given the vertex set, there are $\frac{1}{2}(l - 1)!$ ways of choosing C_0,

$$E(Z_l) = \tfrac{1}{2}(l-1)!\binom{n}{l}\binom{N-l}{M-l}\binom{N}{M}^{-1} \leq \frac{1}{2l}n^l\left(\frac{M}{N}\right)^l.$$

In the inequality above we applied (3) with $a = N, b = M, x = l$ and $y = 0$. Summing over l we find that

$$\sum_{l=3}^{g-1} E(Z_l) \leq \sum_{l=3}^{g-1} \frac{1}{2l}n^l\left\{\frac{\delta n}{N}\right\}^l < \sum_{l=3}^{g-1} (2\delta)^l < n.$$

Hence there is a graph $G \in G(n, \delta n)$ that contains at most $n - 1$ cycles of length at most $g - 1$. Therefore, by omitting one edge from each cycle of length at most $g - 1$, we obtain a graph H of order n and size $M' > (\delta - 1)n$. We know from the simple Exercise 11 of Chapter IV that H contains a graph with minimum degree at least δ. □

The next result solves the second problem posed in the introduction of this chapter.

Theorem 5. *Given natural numbers $k \geq 4$ and $g \geq 4$, there is a k-chromatic graph of girth g (and order $k^{4g} + g$).*

PROOF. Let $\Omega = \mathscr{G}(n, M)$, where $n = k^{4g}$ and $M = k^3 n$. Then, as in the previous proof, we have

$$\sum_{l=3}^{g-1} E(Z_l) < \sum_{l=3}^{g-1} (2k^3)^l < (2k^3)^g < \frac{1}{g}n^{7/8}.$$

Put $f = (2k^3)^g$ and denote by Ω_1 the set of graphs in Ω that contain at most $3f$ cycles of length less than g. Then trivially (or by Chebyshev's inequality, see §4), we have

$$P(\Omega_1) \geq \tfrac{2}{3}.$$

Now we put $p = (n/k) + 1 = k^{4g-1} + 1$ and estimate $P(\Omega_2)$, the measure of the set Ω_2 of graphs G that do not contain a p-set with at most $3f$ edges.

Clearly the graphs in Ω_2 are not $(k-1)$-colourable. Furthermore, even if we omit any $3f$ of their edges, we still cannot colour them with $k-1$ colours. In order to estimate $P(\Omega_2)$ we calculate the expectation of $I_p^{+l} = I_p^{+l}(G)$, the number of p-sets of vertices in which G has exactly l edges. There are $\binom{n}{p}$ ways of choosing a set of p vertices; there are

$$\binom{\binom{p}{2}}{l}$$

ways of choosing l edges joining vertices of such a set; and the remaining $M-l$ edges can be chosen in

$$\binom{N-\binom{p}{2}}{M-l}$$

ways. Thus

$$E(I_p^{+l}) = \binom{n}{p}\binom{\binom{p}{2}}{l}\binom{N-\binom{p}{2}}{M-l}\binom{N}{M}^{-1}.$$

Applying (3) and summing over l we find that

$$\sum_{l=0}^{3f} E(I_p^{+l}) \le \sum_{l=0}^{3f}\left(\frac{en}{p}\right)^p\left(\frac{ep^2}{2}\right)^l \exp\left\{-M\binom{p}{2}\binom{n}{2}^{-1}\right\}$$
$$< k^p e^{p+3f} p^{6f} \exp\{-kn\}.$$

The logarithm of the right hand side is at most

$$n + \frac{n}{3} + n^{7/8} + \tfrac{3}{2}n^{7/8}\log n - kn < -\tfrac{1}{2}n.$$

Consequently

$$\sum_{l=0}^{3f} E(I_p^{+l}) < e^{-n/2} < \tfrac{1}{3}$$

and so

$$P(\Omega_2) > \tfrac{2}{3}.$$

Therefore

$$P(\Omega_1 \cap \Omega_2) \ge \tfrac{1}{3}.$$

Every graph $H \in \Omega_1 \cap \Omega_2$ can be altered in a very simple way to provide a graph with the required properties. Indeed, since $H \in \Omega_1$, it contains at most $3f$ cycles of length less than g. Let F be obtained from H by omitting $3f$ edges of H, at least one from each cycle of length less than g. Then F has girth $g(F) \ge g$. Since $H \in \Omega_2$, this graph F does not contain $p = (n/k) + 1$ independent vertices and so $\chi(F) \ge k$. Let G be the disjoint union of a cycle C^g and a k-chromatic subgraph of F. □

In §1 we promised to sketch a proof of a lower bound for the Ramsey number $R(3, t)$, given by Erdös. The sketch will be rather rough since the details of the proof are somewhat technical and cumbersome.

Theorem 6. *There is a constant $c > 0$ such that*

$$R(3, t) \geq c\left(\frac{t}{\log t}\right)^2$$

for every $t \geq 2$. □

It suffices to show that there is a (large) positive constant A such that for every n there is a graph G^n without triangles and without $\lfloor An^{1/2} \log n \rfloor$ independent vertices. Indeed, it is easily checked that if there is such a constant then $c = \frac{1}{4}A^{-2}$ will do in the theorem.

In order to show the existence of A, we choose a constant $B > 0$ and consider the probability space $\Omega = \mathscr{G}(n, M)$, where $M = \lfloor B^{-1/2}n^{3/2} \rfloor$. Given $G \in \Omega$, let E^* be a minimal set of edges whose removal destroys every triangle in G. If $H = G - E^*$ contains a set W of $p = \lfloor Bn^{1/2} \log n \rfloor$ independent vertices, then every edge of $G[W]$ has to be contained in a triangle of G whose third vertex is outside W. One can show that if B and n are sufficiently large, say $B \geq B_0$ and $n \geq n_0$, then most graphs $G \in \Omega$ do not contain such a set W. Clearly $A = B_0 n_0^2$ has the required properties.

§3 Simple Properties of Almost All Graphs—the Basic Use of Probability

In this section we shall introduce and examine the model $\mathscr{G}(n, P(\text{edge}) = p)$. The graphs in this model are such that the *edges* are chosen *independently* and with the *same probability* p, $0 < p < 1$. Thus $\mathscr{G}(n, P(\text{edge}) = p)$ consists of *all* graphs with a fixed set of n distinguishable vertices, and the probability of a graph with m edges is $p^m q^{N-m}$, where $q = 1 - p$ and, as before, $N = \binom{n}{2}$. Thus q is the probability that a given pair of vertices are not joined by an edge, and N is the maximum possible number of edges. We saw in the preceding section how useful it is to know that most graphs in a model have a certain property. Now we shall go a step further, namely we shall discuss properties shared by almost all graphs. Given a property Q, we shall say that *almost every* (a.e.) *graph* (in a probability space Ω_n consisting of graphs of order n) has property Q if $P(G \in \Omega_n : G \text{ has } Q) \to 1$ as $n \to \infty$. In this section we shall always take $\Omega_n = \mathscr{G}(n, \text{P(edge)} = p)$, where $0 < p < 1$ may depend on n.

Let us assume first that $0 < p < 1$ *is fixed*, that is p is independent of n.

There are many simple properties holding for a.e. graph in $\mathscr{G}(n, P(\text{edge}) = p)$. For instance, if H is an arbitrary fixed graph, then a.e. $G \in \mathscr{G}(n, P(\text{edge}) = p)$ contains H as a spanned subgraph. Indeed, if $|H| = h$, then the probability

that the subgraph of G spanned by a given set of h vertices is isomorphic to H is positive, say $r > 0$. Since $V(G)$ contains $\lfloor n/h \rfloor$ disjoint subsets of h vertices each, the probability that no spanned subgraph of G is isomorphic to H is at most $(1 - r)^{\lfloor n/h \rfloor}$, which tends to 0 as $n \to \infty$. The following result is a strengthened version of this observation.

Theorem 7. *Let $1 \leq h \leq k$ be fixed natural numbers and let $0 < p < 1$ be fixed also. Then in $\mathscr{G}(n, P(edge) = p)$ a.e. graph G is such that for every sequence of k vertices $x_1, x_2, \ldots, x_k$ there exists a vertex x such that $xx_i \in E(G)$ if $1 \leq i \leq h$ and $xx_i \notin E(G)$ if $h < i \leq k$.*

PROOF. Let $x_1, x_2, \ldots, x_k$ be an arbitrary sequence. The probability that a vertex $x \in W = V(G) - \{x_1, \ldots, x_k\}$ has the required properties is $p^h q^{k-h}$. Since for $x, y \in W$, $x \neq y$, the edges xx_i are chosen independently of the edges yx_i, the probability that no suitable vertex x can be found for this particular sequence is $(1 - p^h q^{k-h})^{n-k}$. Since there are $n(n-1)\ldots(n-k+1)$ choices for the sequence $x_1, x_2, \ldots, x_k$, the probability that there is a sequence $x_1, x_2, \ldots, x_k$ for which no suitable x can be found is at most

$$\varepsilon = n^k(1 - p^h q^{k-h})^{n-k}.$$

Clearly $\varepsilon \to 0$ as $n \to \infty$. □

By a result of Gaifman concerning first order sentences, Theorem 7 implies that *for a fixed $0 < p < 1$ every first order sentence about graphs is either true for a.e. graph in $\mathscr{G}(n, P(\text{edge}) = p)$ or is false for a.e. graph.* Though this result looks rather sophisticated, it is in fact weaker than the shallow Theorem 7, for given any first order sentence, Theorem 7 enables us to deduce immediately whether the sentence holds for a.e. graph or it is false for a.e. graph. In particular, each of the following statements concerning the model $\mathscr{G}(n, P(\text{edge}) = p)$ for a fixed p, $0 < p < 1$, is an immediate consequence of Theorem 7.

1. For a fixed integer k, a.e. graph is k-connected.
2. A.e. graph has diameter 2.
3. Given a graph H, a.e. graph G is such that whenever $F_0 \subset G$ is isomorphic to a subgraph F of H, there exists an H_0 isomorphic to H satisfying $F_0 \subset H_0 \subset G$.

Rather naturally, most statements we are interested in are not first order sentences, since they concern large subsets of vertices. "For a given $\varepsilon > 0$, a.e. graph has at least $\frac{1}{2}(p - \varepsilon)n^2$ edges and at most $\frac{1}{2}(p + \varepsilon)n^2$ edges". "Almost no graph can be coloured with $n^{1/2}$ colours". "A.e. graph contains a complete graph of order $(\log n)/(\log(1/p))$". "Given $\varepsilon > 0$, a.e. graph is $\frac{1}{2}(p - \varepsilon)n$-connected". These statements are all true for a fixed p and are easily proved (see Exercises 8–11); however, none of them is a first order sentence.

Now we shall examine the model $\Omega = \mathscr{G}(n, P(\text{edge}) = p)$ under the assumption that $0 < p < 1$ *depends on* n *but* $pn^2 \to \infty$ *and* $(1 - p)n^2 \to \infty$ *as* $n \to \infty$. As before, we put $N = \binom{n}{2}$ and for $M = 0, 1, \ldots, N$ we denote by Ω_M the set of graphs in $\mathscr{G}(n, M)$. Clearly $\Omega = \bigcup_{M=0}^{N} \Omega_M$, and the elements of Ω_M have equal probability both in $\mathscr{G}(n, M)$ and $\mathscr{G}(n, P(\text{edge}) = p)$.

We shall show that the models $\Omega = \mathscr{G}(n, P(\text{edge}) = p)$ and $\mathscr{G}(n, M)$ are very close to each other, provided M is about pN, the expected number of edges of a graph in Ω.

Writing P for the probability in Ω, we see that

$$P(\Omega_M) = \binom{N}{M} p^M q^{N-M}.$$

Hence

$$\frac{P(\Omega_M)}{P(\Omega_{M+1})} = \frac{M+1}{N-M}\frac{q}{p}. \tag{4}$$

This shows that $P(\Omega_N)/P(\Omega_{M+1})$ increases with M, and $P(\Omega_M)$ is maximal for some M satisfying $pN - 1 \le M \le pN + 1$. Furthermore, if $0 < \varepsilon < 1$ and n is sufficiently large then, since $pn^2 \to \infty$ as $n \to \infty$,

$$\frac{P(\Omega_M)}{P(\Omega_{M+1})} < 1 - \varepsilon$$

provided $M < (1 - \varepsilon)pN$, and

$$\frac{P(\Omega_{M+1})}{P(\Omega_M)} < (1 + \varepsilon)^{-1}$$

when $M > (1 + \varepsilon)pN$. Putting $N_\varepsilon = \lfloor (1 + \varepsilon)pN \rfloor$ and $N_{-\varepsilon} = \lceil (1 - \varepsilon)pN \rceil$, we see from these inequalities that a.e. graph G in Ω satisfies $N_{-\varepsilon} \le e(G) \le N_\varepsilon$, that is

$$P\left(\bigcup_{M=N_{-\varepsilon}}^{N\varepsilon} \Omega_M\right) \to 1 \quad \text{as } n \to \infty. \tag{5}$$

Another consequence of (4) is that there is an $\eta > 0$ (in fact, any $0 < \eta < \frac{1}{2}$ would do) such that

$$P\left(\bigcup_{M=0}^{pN} \Omega_M\right) > \eta \tag{6}$$

if n is sufficiently large. Now (5) and (6) imply that if $\Omega^* \subset \Omega$ is such that $P(\Omega^*) \to 1$ as $n \to \infty$, then for any $\varepsilon > 0$ there are M_1 and M_2, such that $(1 - \varepsilon)pN \le M_1 \le pN \le M_2 \le (1 + \varepsilon)pN$ and

$$\frac{|\Omega_{M_i} \cap \Omega^*|}{|\Omega_{M_i}|} \to 1 \quad \text{as } n \to \infty \qquad (i = 1, 2,). \tag{7}$$

We call a set $\Omega^* \subset \Omega$ *convex* if $G \in \Omega^*$ whenever $G_1 \subset G \subset G_2$ and $G_1, G_2 \in \Omega^*$; a convex property of graphs is defined analogously. It is easily seen that for a convex set Ω^* relation (7) implies that

$$\frac{|\Omega_M \cap \Omega^*|}{|\Omega_M|} \to 1 \quad \text{as } n \to \infty, \tag{7'}$$

whenever $M_1 \leq M \leq M_2$, in particular, if $M = \lfloor pN \rfloor$. Let us restate the assertions above as a theorem about the connection between the models $\mathscr{G}(n, P(\text{edge}) = p)$ and $\mathscr{G}(n, M)$.

Theorem 8. *Let $0 < p = p(n) < 1$ be such that $pn^2 \to \infty$ and $(1 - p)n^2 \to \infty$ as $n \to \infty$, and let Q be a property of graphs.*

(i) *Suppose $\varepsilon > 0$ is fixed and, if $(1 - \varepsilon)pN < M < (1 + \varepsilon)pN$, then a.e. graph in $\mathscr{G}(n, M)$ has Q. Then a.e. graph in $\mathscr{G}(n, P(edge) = p)$ has Q.*

(ii) *If Q is a convex property and a.e. graph in $\mathscr{G}(n, P(edge) = p)$ has Q, then a.e. graph in $\mathscr{G}(n, \lfloor pN \rfloor)$ has Q.* □

§4 Almost Determined Variables—the Use of the Variance

If $X = X(G)$ is a non-negative variable on $\Omega = \mathscr{G}(n, M)$ or $\Omega = \mathscr{G}(n, P(\text{edge}) = p)$, and the expectation of X is at most a, then

$$P(X \leq ta) \geq \frac{t - 1}{t} \quad \text{for } t > 1.$$

Thus if the expectation of X is very small then X is small for most graphs. This simple fact was used over and over again in the first two sections. However, if we want to show that X is *large* or *non-zero* for almost every graph in Ω then the expected value itself very rarely can help us, so we have to try a slightly more sophisticated attack. We usually turn to the *variance* for help. Recall that if $\mu = E(X)$ is the expectation of X then

$$\operatorname{Var}(X) = \sigma^2(X) = E((X - \mu)^2) = E(X^2) - \mu^2$$

is the *variance* of X and $\sigma = \sigma(X) > 0$ is the *standard deviation*. Chebyshev's inequality, which is immediate from first principles, states that if $t > 0$ then

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}.$$

In particular, if $0 < t < \mu$ then

$$P(X = 0) \le P(|X - \mu| \ge t) \le \frac{\sigma^2}{t^2},$$

so

$$P(X = 0) \le \frac{\sigma^2}{\mu^2}. \tag{8}$$

In the examples we consider $X = X(G)$ is always the number of *spanned* subgraphs of G contained in a family $\mathscr{F} = \{F_1, F_2, \ldots\}$. Here the family $\mathscr{F}$ may depend on n and every graph $F \in \mathscr{F}$ has the same labelled vertex set as G. Then clearly

$$E(X^2) = \sum_{(F', F'')} P(G \text{ contains } F' \text{ and } F''), \tag{9}$$

where the summation is over all *ordered* pairs (F', F''), $F', F'' \in \mathscr{F}$.

In this section we shall consider the space $\Omega = \mathscr{G}(n, P(\text{edge}) = p)$, where $0 < p < 1$ is fixed; we know that this space is near to $\mathscr{G}(n, M)$, where $M = \lfloor pn^2/2 \rfloor$. As before, we write $q = 1 - p$ for the probability that two given vertices are not adjacent. Furthermore, we use Landau's notation $O(f(n))$ for a term that, when divided by $f(n)$, remains bounded as $n \to \infty$. Similarly, $o(f(n))$ denotes a term that, when divided by $f(n)$, tends to 0 as $n \to \infty$. Thus $O(1)$ is a bounded term and $o(1)$ is a term tending to 0.

Our aim is to make use of the variance to show that certain graph invariants are almost determined in our model $\mathscr{G}(n, P(\text{edge}) = p)$. The first theorem concerns the maximal degree. In its proof we shall make use of a special case of the classical De Moivre–Laplace theorem concerning the approximation of the binomial distribution by the normal distribution. Let $0 < c = c(n) = O(1)$ and put $d(c) = \lfloor pn + c(pqn \log n)^{1/2} \rfloor$. Suppose $x(c) = c(\log n)^{1/2} \to \infty$. Then

$$\begin{aligned}\sum_{k=d(c)}^{n} \binom{n}{k} p^k q^{n-k} &= (1 + o(1)) \frac{1}{\sqrt{2\pi}} \int_{x(c)}^{\infty} e^{-u^2/2}\, du \\ &= (1 + o(1))(2\pi c^2 \log n)^{-1/2} n^{-c^2/2}. \end{aligned}\tag{10}$$

Furthermore

$$\begin{aligned}\sum_{k=d(c)}^{n} \left\{\binom{n}{k} p^k q^{n-k}\right\}^2 &= (1 + o(1))(2\pi pqn)^{-1/2} \int_{x(c)}^{\infty} e^{-u^2}\, du \\ &= (1 + o(1))(8\pi pqc^2 n \log n)^{-1/2} n^{-c^2} \\ &= o(n^{-c^2 - 1/2}). \end{aligned}\tag{11}$$

Theorem 9. *If $0 < p < 1$ is fixed then in $\Omega = \mathscr{G}(n, P(edge) = p)$ the maximal degree of almost every graph is*

$$pn + (2pqn \log n)^{1/2} + o(n \log n)^{1/2}.$$

Proof. Let $c > 0$ and denote by $X_c = X_c(G)$ the number of vertices of degree at least $d(c) = \lfloor pn + c(pqn \log n)^{1/2} \rfloor$. Then, by (10),

$$\begin{aligned}\mu_c = E(X_c) &= n \sum_{k=d(c)}^{n-1} \binom{n-1}{k} p^k q^{n-1-k} \\ &= (1 + o(1))(2\pi c^2 \log n)^{-1/2} n^{1-(c^2/2)}\end{aligned}$$

and

$$\begin{aligned}E(X_c^2) &\le E(X_c) \\ &\quad + n(n-1) \sum_{k_1=d(c)}^{n-1} \sum_{k_2=d(c)}^{n-1} \binom{n-2}{k_1-1}\binom{n-2}{k_2-1} p^{k_1+k_2-2} q^{2n-2-k_1-k_2} \\ &= \mu_c + (1 + o(1))\mu_c^2.\end{aligned}$$

In the estimation of $E(X_c^2)$ the term

$$\binom{n-2}{k_1-1}\binom{n-2}{k_2-1} p^{k_1+k_2-2} q^{2n-2-k_1-k_2}$$

is the probability that, given vertices $a, b \in G$, $a \neq b$, the vertex a is joined to $k_1 - 1$ vertices in $V(G) - \{a, b\}$ and b is joined to $k_2 - 1$ vertices in $V(G) - \{a, b\}$.

Now if $c = \sqrt{2} - \varepsilon$, where $0 < \varepsilon < \frac{1}{2}$, then $\mu_c \ge (1 + o(1))n^{\varepsilon}$, so $E(X_c^2) = (1 + o(1))\mu_c^2$ and $\sigma^2(X_c) = o(\mu_c^2)$. Therefore, by (8),

$$P(X_c = 0) = o(1).$$

Thus G almost surely has a vertex of degree at least

$$pn + (\sqrt{2} - \varepsilon)(pqn \log n)^{1/2}.$$

On the other hand, if $c = \sqrt{2} + \varepsilon$, $\varepsilon > 0$, then $\mu_c = o(n^{-\varepsilon})$ so

$$P(X_c \ge 1) = o(n^{-\varepsilon}). \qquad \square$$

The result above has an interesting consequence.

Corollary 10. *If $0 < p < 1$ and $c > \sqrt{(3/2)}$ then in $\mathscr{G}(n, P(edge) = p)$ almost no graph has two vertices of equal degree whose degrees are at least*

$$d(c) \doteq \lfloor pn + c(pqn \log n)^{1/2} \rfloor.$$

Proof. The probability of the existence of two such vertices is at most

$$n^2 \sum_{k=d(c)}^{n-1} \left\{ \binom{n-2}{k-1} p^{k-1} q^{n-1-k} \right\}^2$$

and by (11) this is $o(1)$. □

Combining Theorem 9 with Corollary 10 we see that a.e. graph has a unique vertex with maximum degree.

In the next theorem we again consider the case when p varies with n. We say that a function $t(n)$ is a *threshold function* for a property Q of graphs if

$$\lim_{n \to \infty} P_{n,p}(Q) = \begin{cases} 0 & \text{if } p/t(n) \to 0, \\ 1 & \text{if } p/t(n) \to \infty. \end{cases}$$

Here $P_{n,p}(Q)$ denotes the probability that in $\mathscr{G}(n, P(\text{edge}) = p)$ a graph has property Q. The existence of a threshold function means that in the evolution of a random graph, i.e. in the process a graph acquires more and more edges, the property emerges rather abruptly. In this section we prove a basic result about threshold functions. The next section will be devoted to the beautiful result that the property of being Hamiltonian has a sharp threshold function.

Theorem 11. *Let $k \geq 2$, $k - 1 \leq l \leq \binom{k}{2}$ and let $F = G(k, l)$ be a graph with average degree $(2l/k)$ at least as large as that of any of its subgraphs. Then $n^{-k/l}$ is a threshold function for F, that is if $pn^{k/l} \to 0$ then in $\mathscr{G}(n, P(edge) = p)$ almost no graph contains F and if $pn^{k/l} \to \infty$ then almost every graph contains F.*

Proof. Let $p = \gamma n^{-k/l}$, $0 < \gamma < n^{k/l}$, and denote by $X = X(G)$ the number of copies of F contained in $G \in \mathscr{G}(n, P(\text{edge}) = p)$. Denote by k_F the number of graphs with a fixed set of k labelled vertices that are isomorphic to F. Clearly $k_F \leq k!$. Then

$$\mu = \mu_\gamma = E(X) = \binom{n}{k} k_F p^l (1-p)^{\binom{k}{2} - l} \leq n^k (\gamma^l n^{-k}) = \gamma^l,$$

so $E(X) \to 0$ as $\gamma \to 0$, showing the first assertion.

Now let us estimate the variance of X when γ is large. Note that there is a constant $c_1 > 0$ such that

$$\mu_\gamma \geq c_1 \gamma^l \quad \text{for every } \gamma. \tag{12}$$

According to (9), we have to estimate the probability that G contains two fixed copies of F, say F' and F''. Put

$$A_s = \sum_s P(G \text{ contains } F' \text{ and } F''),$$

where $\sum_s$ means that the summation is over all pairs (F', F'') with s vertices in common. Clearly

$$A_0 < \mu^2.$$

Furthermore, in a set of s vertices F' has $t \le (l/k)s$ edges. Hence, counting first the choices for F' and then for F'' with $s \ge 1$ common vertices with F', we find that for some constants c_2 and c_3,

$$\begin{aligned} \frac{A_s}{\mu} &\le \sum_{t \le ls/k} \binom{k}{s}\binom{n-k}{k-s} k! p^{l-t} q^{\binom{k}{2}-l+t} \\ &\le \sum_{t \le ls/k} c_2 n^{k-s} (\gamma n^{-k/l})^{l-t} \\ &\le c_2 m^{-s}\gamma^l + c_3 \gamma^{l-1}. \end{aligned}$$

Here in the last step we separated the term with $t = 0$ from the rest. Consequently, making use of (12), we find that

$$\frac{E(X^2)}{\mu^2} = \frac{\sum_0^k A_s}{\mu^2} \le 1 + c_4 \gamma^{-1}$$

for some constant c_4. Therefore, by (8),

$$P(X = 0) \le \frac{\sigma^2}{\mu^2} \le c_4 \gamma^{-1}$$

so $P(X = 0) \to 0$ as $\gamma \to \infty$. □

The most striking example of a graphic invariant being almost determined in a random graph is that of the *clique number*, the maximal order of a complete subgraph. It turns out that for a fixed p, $0 < p < 1$, the clique number of almost every graph in $\mathcal{G}(n, P(\text{edge}) = p)$ is one of *two* possible values. In fact, for most values of n (in a well defined sense) the clique number of almost every graph is just a function of p and n. We shall confine ourselves to proving a simple result in this direction. As in Theorem 1, denote by $X = X(G)$ the number of K^r subgraphs in $G \in \mathcal{G}(n, P(\text{edge}) = p)$. Then

$$E(X_r) = \binom{n}{r} p^{\binom{r}{2}},$$

since we have $\binom{n}{r}$ choices for the vertex set of a K^r and, having chosen the vertex set, we have no choice for the $\binom{r}{2}$ edges. Let $d = d(n, p)$ be the positive *real* number for which

$$\binom{n}{d} p^{\binom{d}{2}} = 1.$$

For simplicity we put $b = 1/p$. It is easily checked with the aid of Stirling's formula that

$$b^{d/2} < n$$

and

$$d = \frac{2 \log n}{\log b} + O(\log \log n). \tag{6}$$

Theorem 12. *Let $0 < p < 1$ be fixed. Then the clique number of almost every graph $G \in \mathcal{G}(n, P(edge) = p)$ is $\lfloor d \rfloor$ or $\lceil d \rceil$.*

PROOF. The assertion is equivalent to the following:

$$P(X_r > 0) \to 0 \quad \text{if } r \geq d + 1,$$
$$P(X_r > 0) \to 1 \quad \text{if } r \leq d - 1.$$

Now if $r \geq d + 1$, then by (6),

$$E(X_r) = \binom{n}{r} p^{\binom{r}{2}} \leq \frac{n}{r} p^r \binom{n}{r-1} p^{\binom{r-1}{2}} \leq \frac{n}{r} p^r \to 0,$$

so the first assertion is clear.

Let now $r \leq d - 1$. Let us use (9) to calculate the second moment of X_r, summing separately over pairs of K^r subgraphs with exactly l vertices in common:

$$E(X_r^2) = \sum_{l=0}^{r} \binom{n}{r}\binom{r}{l}\binom{n-r}{r-l} p^{2\binom{r}{2}} = \binom{n}{r}^2 p^{2\binom{r}{2}-\binom{l}{2}}.$$

Since

$$\mu_r^2 = E(X_r)^2 = \sum_{l=0}^{r} \binom{n}{r}\binom{n}{l}\binom{n-r}{r-l} p^{2\binom{r}{2}} = \binom{n}{r}^2 p^{2\binom{r}{2}}$$

with $\sigma_r = \sigma(X_r)$ we have

$$\frac{\sigma_r^2}{\mu_r^2} = \frac{E(X_r^2)}{\mu_r^2} - 1 \leq \sum_{l=2}^{r} \frac{\binom{r}{l}\binom{n-r}{r-l}}{\binom{n}{r}} p^{\binom{l}{2}}$$

$$\leq \sum_{l=2}^{r} r^{2l} n^{-l} b^{(l-1)l/2}.$$

Recall now that $b^{d/2} < n$ so $b^{(l-1)l/2} < n^{l-1}$. Hence

$$\frac{\sigma_r^2}{\mu_r^2} \le r^{2r+1} n^{-1} \to 0 \quad \text{as } n \to \infty,$$

proving the second assertion. □

§5 Hamilton Cycles—the Use of Graph Theoretic Tools

In the proofs so far we always adopted a more or less head-on attack. We hardly needed more from graph theory than the definitions of the concepts involved, the emphasis was on the use of probability theory. This section is devoted to a beautiful result of Pósa, the proof of which is based on a non-trivial result in graph theory. Of course, in the ideal use of probabilistic methods in graph theory we would have a mixture of all the ideas presented in the four sections. Thus we would prepare the ground by using non-trivial graph theoretic results and would apply probability theory to get information about graphs in a probability space tailor made for the problem. We could then select an appropriate graph which we would afterwards alter with the aid of powerful graph theoretic tools.

It is easily seen that if $\varepsilon > 0$ then almost all graphs of order n and size $M = \lfloor(\frac{1}{2} + \varepsilon)n \log n\rfloor$ are connected and one can also show that almost no graph of order n and size $M = \lfloor(\frac{1}{2} - \varepsilon)n \log n\rfloor$ is connected (Exercise 13). In particular, almost no graph of order n and size $M = \lfloor(\frac{1}{2} - \varepsilon)n \log n\rfloor$ contains a Hamilton cycle. It is fascinating that roughly the same number of edges as are needed to ensure connectivity already ensure the existence of a Hamilton cycle. Our next aim is to prove this beautiful theorem of Pósa.

The basis of the proof of this result is Theorem 15 of Chapter IV. Let S be a longest x_0-path in a graph H and write L for the set of endvertices of the transforms of S. Denote by N the set of neighbours of vertices of L *on* S and put $\mathscr{R} = V(H) - L \cup N$. Then Theorem 15 of Chapter IV states that H has no $L - \mathscr{R}$ edge. All we shall need from this is that if $|L| = l \le |H|/3$ then there are disjoint sets of size l and $|H| - 3l + 1$, that are joined by no edge of H.

For the sake of convenience we shall work with the space

$$\mathscr{G}(n, P(\text{edge}) = p), \quad p = \frac{c \log n}{n}.$$

We start with a simple lemma in the vein of Theorem 4. Denote by D_t the number of pairs (X, Y) of disjoint subsets of V such that $|X| = t$, $|Y| = n - 1 - 3t$, and G has no $X - Y$ edge.

Lemma 13. *Let $c > 3$ and $0 < \gamma < \frac{1}{3}$ be constants and let $p = (c \log n)/n$. Then in $\mathcal{G}(n, P(edge) = p)$ we have*

$$P(D_t > 0 \text{ for some } t, 1 \le t \le \gamma n) = O(n^{3-c}).$$

PROOF. Put $\beta = (c - 3)/4c$. Clearly

$$\begin{aligned}\sum_{t=1}^{\lfloor \gamma n \rfloor} E(D_t) &= \sum_{t=1}^{\lfloor \gamma n \rfloor} \binom{n}{t}\binom{n-t}{n-3t}(1-p)^{t(n-3t)} \\ &\le n\binom{n-1}{2}(1-p)^{n-1} + \sum_{t=2}^{\lfloor \beta n \rfloor} \frac{1}{t!} n^{3t}(1-p)^{t(n-3t)} \\ &\quad + \sum_{\lfloor \beta n \rfloor + 1}^{\lfloor \gamma n \rfloor} 2^{2n}(1-p)^{t(n-3t)}.\end{aligned}$$

Now since $(1 - p)^n < n^{-c}$, we have

$$n^3(1-p)^{n-3} < (1-p)^3 n^{3-c},$$

$$n^{3t}(1-p)^{t(n-3t)} < n^{t(3-(c(n-3t)/n))} \le n^{3-c} \quad \text{if } 2 \le t \le \beta n$$

and

$$2^{2n}(1-p)^{t(n-3t)} < n^{(2n/(\log n)) - ((n-3t)t/n)} = O(n^{-\beta(1-3\gamma)n}) \quad \text{if } \beta n \le t \le \gamma n.$$

Consequently

$$\sum_{t=1}^{\lfloor \gamma n \rfloor} E(D_t) = O(n^{3-c}),$$

implying the assertion of the lemma. □

Theorem 14. *Let $p = (c \log n)/n$ and consider the space $\mathcal{G}(n, P(edge) = p)$. If $c > 3$ and x and y are arbitrary vertices, then almost every graph contains a Hamilton path from x to y. If $c > 9$ then almost every graph is Hamiltonian connected: every pair of distinct vertices is joined by a Hamilton path.*

PROOF. Choose $\gamma < \frac{1}{3}$ in such a way that $c\gamma > 3$ if $c > 9$ and $c\gamma > 1$ if $c > 3$.

Let us introduce the following notation for certain events in $\mathcal{G}(n, P(\text{edge}) = p)$, whose general element is denoted by G.

$D = \{D_t = 0 \text{ for every } t, 1 \le t \le \lfloor \gamma n \rfloor\}$,

$E(W, x) = \{G[W]$ has a path of maximal length, whose endvertex is joined to $x\}$,

$E(W, x|w) = \{G[W]$ has a w-path of maximal length among the w-paths, whose endvertex is joined to $x\}$,

$F(x) = \{$every path of maximal length contains $x\}$,

$H(W) = \{G[W]$ has a Hamilton path$\}$,

$H(x, y) = \{G$ has a Hamilton x-y path$\}$,

$HC = \{G$ is Hamiltonian connected$\}$.

The complement of an event A is $\bar{A}$.

Note that by Lemma 13 we have

$$P(\bar{D}) = 1 - P(D) = O(n^{c-3}).$$

Let $|W| = n - 2$ or $n - 1$ and let us estimate $P(D \cap \bar{E}(W, x))$, where $x \notin W$. Let $G \in D \cap \bar{E}(W, x)$ and consider a path $S = x_0 x_1 \dots x_k$ of maximal length in $G[W]$. (By introducing an ordering in W, we can easily achieve that S is determined by $G[W]$.) Let $L = L(G[W])$ be the set of endvertices of the transforms of the x_0-path S and let M be as in Theorem 15 of Chapter IV (applied to $G[W]$). Recall that $|M| \geq |W| + 1 - 3|L|$ and there is no $L - M$ edge, so no $L - M \cup \{x\}$ edge either. Since $G \in D$ and $|M \cup \{x\}| \geq n - 3|L|$, we find that $|L| \geq \gamma n$. As L was determined *independently* of the edges incident with x, we have

$$\begin{aligned} P(D \cap \bar{E}(W, x)) &\leq P(G \in D \quad \text{and } x \text{ is not joined to } L(G[W])) \\ &\leq (1 - p)^{\gamma n} < n^{-c\gamma}. \end{aligned}$$

Exactly the same proof implies that

$$P(D \cap \bar{E}(W, x|w)) < n^{-c\gamma}$$

provided $|W| = n - 2$ or $n - 1$, $w \in W$ and $x \notin W$.

Note now that $\bar{F}(x) \subset \bar{E}(V - \{x\}, x)$, so

$$\begin{aligned} P(\bar{H}(V)) &= P\Big(\bigcup_{x \in V} \bar{F}(x)\Big) \leq P\Big(D \cap \bigcup_{x \in V} \bar{F}(x)\Big) + P(\bar{D}) \\ &\leq \sum_{x \in V} P(D \cap \bar{F}(x)) + P(\bar{D}) \leq nP(D \cap \bar{E}(V - \{x\}, x)) + P(\bar{D}) \\ &\leq n^{1-c\gamma} + O(n^{3-c}). \end{aligned}$$

This proves that if $c > 3$ then almost every graph has a Hamilton path.

Now let x and y be distinct vertices and put $W = V - \{x, y\}$. By the first part

$$P(\bar{H}(W)) \leq 2n^{1-c\gamma} + O(n^{3-c}).$$

Since

$$H(x, y) \subset H(W) \cap E(W, y) \cap E(W, x|y),$$

we have

$$\begin{aligned} P(\bar{H}(x, y)) &\leq P(\bar{H}(W)) + P(D \cap \bar{E}(W, y)) + P(D \cap \bar{E}(W, x|y) + P(\bar{D}) \\ &\leq 2n^{1-c\gamma} + 2n^{-c\gamma} + O(N^{c-3}). \end{aligned}$$

Therefore, if $c > 3$ then almost every graph contains a Hamilton path from x to y.

Finally, as there are $\binom{n}{2}$ choices for an unordered pair (x, y), $x \neq y$,

$$P(\overline{HC}) \leq \sum_{x=y} P(\overline{H}(x, y)) \leq n^{3-c\gamma} + O(n^{2-c\gamma}).$$

Thus if $c > 9$ then almost every graph is Hamiltonian connected. □

Since every Hamiltonian connected graph is Hamiltonian (contains a Hamilton cycle), by Theorem 8 we have, in particular, that if $c > 9$ then almost every graph in $\mathcal{G}(n, \lfloor (c \log n)/n \rfloor)$ is Hamiltonian. In fact, one can see that the essential part of the proof gives this result with $c > 6$ as well. Independently of Pósa, Korshunov proved a sharper and essentially best possible form of Theorem 14: the assertion holds for every $c > 1$.

Exercises

1. Let $\vec{G} = (V, \vec{E})$ be a directed graph with m edges and without loops. Prove that V can be partitioned into sets V_1 and V_2 such that $\vec{G}$ contains more than $m/4$ edges from V_1 to V_2.

2. Given an integer $k \geq 2$ and a graph G, denote by $p^{(k)}(G)$ the minimal number of edges of G whose *omission* results in a k-partite graph. Prove that if $G = G(n, m)$ then

$$p^{(k)}(G) \leq m\left\{l\binom{r+1}{2} + (k-l)\binom{r}{2}\right\}\binom{n}{2}^{-1},$$

where $n = kr + l$, $0 \leq l < k$. Show furthermore that

$$m - kp^{(k)}(G) \geq \frac{k-1}{n-1}m.$$

[*Hint.* Consider all graphs of the form $H = lK^{r+1} \cup (k-l)K^r$ with vertex set $V(G)$. Note that the expected number of edges common to G and H is $e(G)e(H)/e(K^n)$.]

3. Prove that there is a tournament of order n (see Exercise 12 of Chapter I) that contains at least $n!2^{-n+1}$ directed Hamiltonian paths.

4. Show that there is a graph of order n and size $\lfloor \frac{1}{2}(1 - (1/s!t!))n^{2-((s+t-2)/(st-1))} \rfloor$ that does not contain a $K(s, t)$.

5. Given $2 \leq s \leq n$, let d be the maximal integer for which there is a $G_3(n, n, n)$ without a $K_3(s, s, s)$, in which every vertex is joined by at least d edges to each of the other two classes. Prove a lower bound for d.

6. Use Theorem 3 to prove that if $r > 2$, $0 < \varepsilon < \frac{1}{2}(r-1)^{-2}$ and

$$d_r^* > -\frac{2}{\log(2(r-1)^2\varepsilon)}$$

then for every sufficiently large n there is a graph $G(n, m)$ not containing a $K_r(t)$, where $m \geq \{(r-2)/2(r-1) + \varepsilon\}n^2$ and $t = \lfloor d_r^* \log n \rfloor$. (Note that this shows that Theorem 20 of Chapter IV is essentially the best possible.)

7. Show that a given vertex has degree 1 in about $2/e^2$ of the graphs in $\mathcal{G}(n, n)$.

In Exercises 8–13 the model $\mathcal{G}(n, P(\text{edge}) = p)$ is used and $0 < p < 1$ is assumed to be fixed. The same model is used in Exercises 14–19 but p varies with n as indicated.

8.$^-$ Show that for $\varepsilon > 0$ a.e. graph has at least $\frac{1}{2}(p - \varepsilon)n^2$ edges and at most $\frac{1}{2}(p + \varepsilon)n^2$ edges.

9.$^-$ Give a lower bound on the chromatic number of a.e. graph.

10. Show that a.e. graph contains a complete graph of order at least $\log n/\log(1/p)$.

11. Prove that a.e. graph G satisfies.

$$\delta(G) = \lambda(G) = \kappa(G) = pn - (2pqn \log n)^{1/2} + o(n \log n)^{1/2},$$

where $q = 1 - p$.

12. Estimate the maximal value of t for which a.e. graph contains a spanned $K^{t,t} = K(t, t)$. Estimate the corresponding value for $K_r(t) = K(t, \ldots, t)$.

13. Let $0 < c < 1$. Prove that a.e. graph has the property that for every set W of $k = \lfloor c \log_2 n \rfloor$ vertices there is a vertex x_Z for each subset Z of W such that x_Z is joined to each vertex in Z and to none in $W - Z$. Note that for $c = 1$ it is impossible to find even a set of 2^k vertices disjoint from W. [*Hint*. Refine the proof of Theorem 7.]

14. Prove that $(\log n)/n$ is a sharp threshold function for connectivity, in the sense that if $\varepsilon > 0$ and $p = ((1 - \varepsilon)\log n)/n$ then almost no graph is connected while if $\varepsilon > 0$ and $p = ((1 + \varepsilon)\log n)/n$ then a.e. graph is connected.

15.$^+$ Sharpen the previous result as follows. If $p = ((\log n)/n) + (2x/n)$ then the probability that G is connected is $e^{-e^{-2x}}$. [Show first that a.e. G consists of a component and isolated vertices.]

16. Let $p = ((\log n)/n) + (C(n)/n)$, where $C(n) \to \infty$ arbitrarily slowly. Prove that a.e. G contains a 1-factor. [Use Tutte's theorem, Theorem 12 of Chapter III, ignoring the parity of the components.]

17. Show that $1/n$ is a threshold function for F_1 in Figure VII.1, that is if $pn \to 0$ then almost no graph contains F_1 and if $pn \to \infty$ then a.e. graph does.

18. What is the threshold function for F_2 in Figure VII.1?

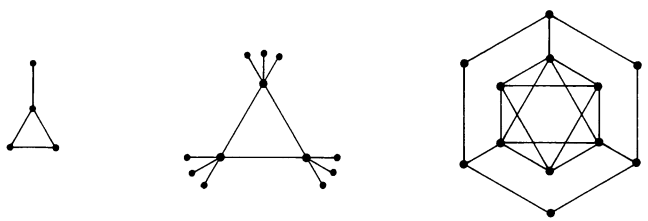


Figure VII.1. The graphs F_1, F_2 and F_3.

19. Let $\varepsilon > 0$. Prove that if $p = n^{-(1/2)-\varepsilon}$ then almost no graph contains F_3 in Figure VII.1 but if $p = n^{-(1/2)+\varepsilon}$ then a.e. graph does. [Find a suitable graph F_3^* that has average degree $2 + \varepsilon$.]

20.⁻ Consider random directed graphs in which all edges are chosen independently and with the same probability p. Prove that there is a constant c such that if $p = c((\log n)/n)^{1/2}$ then a.e. directed graph contains a directed Hamilton cycle. [*Hint*. What is the probability that a graph contains both edges $\vec{ab}$ and $\vec{ba}$? Apply Theorem 14 to the random graph formed by the double edges.]

21.⁻ Note that the suggested solution of Exercise 20 gives two directed Hamilton cycles with the same underlying (non-directed) edge set. Show that with $p = (1 - \varepsilon)((\log n)/n)^{1/2}$ almost no directed graph contains such a pair of Hamilton cycles.

22. Show that there are at least $(2^N/n!) + o(2^N/n!)$ non-isomorphic graphs of order n. [*Hint*. Show that a.e. graph in $\mathscr{G}(n, P(\text{edge}) = \frac{1}{2})$ has trivial automorphism group; for the automorphism group see p. 157.]

Notes

Perhaps the first combinatorial result proved by probabilistic methods is the assertion of Exercise 3, proved by T. Szele in Combinatorial investigations concerning directed complete graphs (in Hungarian), *Mat. Fiz. Lapok* **50** (1943) 223–256; for a German translation see Kombinatorische Untersuchungen über gerichtete vollständige Graphen, *Publ. Math. Debrecen* **13** (1966) 145–168. However, the first papers that gave real impetus to the use of random graphs are two papers of P. Erdös: Graph theory and probability, *Canad. J. Math.* **11** (1959) 34–38, contains Theorem 6 and Graph theory and probability II, *Canad. J. Math.* **13** (1961) 346–352, contains the lower bound on R(3, t) given in Theorem 6.

The result about first order sentences which we mentioned after Theorem 7 is due to R. Fagin, Probabilities on finite models, *J. Symb. Logic* **41** (1976) 50–58.

The fundamental paper on the growth of random graphs is P. Erdös and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960) 17–61. This paper contains a detailed discussion of sparse random graphs, covering amongst other things the distribution of their components and the occurrence of small subgraphs (Theorem 11).

The sharpest results in the direction of Theorem 12 are in B. Bollobás and P. Erdös, Cliques in random graphs, *Math. Proc. Cambridge Phil. Soc.* **80** (1976) 419–427. Pósa's theorem (Theorem 14) is in L. Pósa, *Discrete Math.* **14** (1976) 359–364, its sharper form is in A. D. Korshunov, Solution of a problem of Erdös and Rényi, on Hamilton cycles in nonoriented graphs, *Soviet Mat. Doklady* **17** (1976) 760–764.

For standard results of probability theory, including Chebyshev's inequality and the approximation of the binomial distribution by the Poisson

distribution the reader is referred to W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1 3rd ed., Wiley, New York, 1974, and A. Rényi, *Foundations of Probability*, Holden-Day, San Francisco, 1970, especially to the sections on the De Moivre–Laplace formula.

Random graphs and other combinatorial structures are investigated in depth in the book by P. Erdös and J. Spencer, *Probabilistic Methods in Combinatorics*, Academic Press, New York and London, 1974.

CHAPTER VIII
Graphs and Groups

In this chapter we examine some interactions between graphs and groups. Some problems concerning groups are best attacked by using graphs. Although in this context a graph is hardly more than a visual or computational aid, its use does make the presentation clearer and the problems more manageable. The methods are useful both in theory and in practice: they help us to prove general results about groups and particular results about individual groups. The first section, about Cayley and Schreier diagrams, illustrates well both these aspects. It also contains an informal account of group presentations.

If each member of a class of graphs is known to have particularly pleasant symmetry properties, in other words if it has a large automorphism group, then we may often make use of this to obtain additional information about the class. In some cases this extra information enables us to determine the class almost completely. The second section shows how we can do this using the properties of the adjacency matrix.

Some classes of labelled graphs are easily enumerated, as we shall see in §3. Other enumeration problems, such as counting isomorphism classes of graphs, lead us to the study of orbits of permutation groups. The highlight of the section is the fundamental theorem enumerating such orbits, namely Pólya's theorem.

§1 Cayley and Schreier Diagrams

Let A be a group generated by $a, b, \ldots$. The *graph* of A, also called its *Cayley diagram*, with respect to these generators is a *directed multigraph whose edges are coloured with the generators*: there is an edge *from x to y coloured*

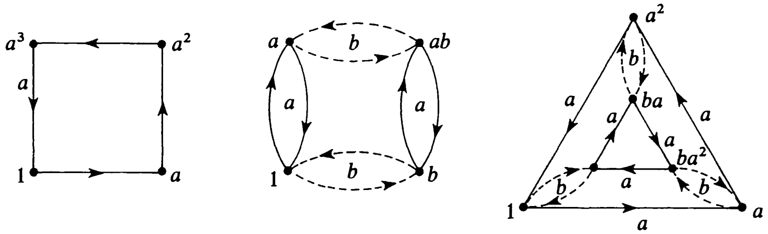


Figure VIII.1. The Cayley diagrams of (i) the cyclic group C_4 generated by a, (ii) the Klein four-group with generators a, b and (iii) the symmetric group S_3 with generators $a = (123)$ and $b = (12)$.

with a generator g iff $xg = y$. To illustrate this concept in Figure VIII.1, we show the Cayley diagrams of three small groups.

A Cayley diagram is *regular*, and so is its colouring, in the following sense: for each vertex x and each generator (colour) g there is exactly one edge of colour g starting at x and exactly one edge of colour g ending at x. Furthermore, at most one edge goes from x to another vertex y. If we know the Cayley diagram of a group then we can easily answer questions in terms of the generators. What is the element aba^2b in S_3? It is the end of the directed walk starting at 1 whose first edge has colour a, the second has colour b, the third a, the fourth a and, finally, the fifth b. By following this walk in the third picture in Figure 1, we find that $aba^2b = a^2$. In general, two elements expressed as products of some generators are equal iff the corresponding walks starting at 1 end at the same vertex.

The Schreier diagram is a slight extension of the Cayley diagram. This time we have a group A, a set S of elements of A and a subgroup B of A. The *Schreier diagram of A mod B* describes the effect of the elements of S on the right cosets of B: it is a directed multigraph whose vertices are the right cosets of B, in which an edge of colour $s \in S$ goes *from* a coset H *to* a coset K iff $Hs = K$. (Thus a Cayley diagram is a Schreier diagram mod $B = \{1\}$.) In most cases S is chosen to be a set of generators or a set which together with B generates A. Instead of giving a separate illustration, note that if A is the symmetric group on $\{1, 2, 3, 4\}$, B is the subgroup of elements fixing 1 and $S = \{a\}$ with $a = (1234)$ then the Schreier diagram is exactly the first picture in Figure VIII.1, that is the Cayley diagram of C_4. Once again we note that for each vertex H and each colour $g \in S$ exactly one edge coloured g starts at H and exactly one edge coloured g ends at H. However, some of the edges may be *loops*, that is they may start amd end at the same coset. Furthermore, there may be many edges of different colours joining two vertices.

Group diagrams do not tell us anything about groups that cannot be expressed algebraically. However, the disadvantage of the algebraic approach is that many lines of print are needed to express what is conveyed almost instantaneously by a single diagram. These diagrams are especially

helpful when we have a concrete problem to be solved by hand, and moreover the purely mechanical techniques involved are ideal for direct use on a computer. Since the advent of fast electronic computers many otherwise hopeless problems have been solved in this way.

Group diagrams are particularly useful when attacking problems concerning groups given by means of their *presentations*. For the convenience of the reader we recall the basic facts about group presentations. We aim throughout for an intuitive description, rather than a rigorous treatment; the interested reader may fill in the details himself or turn to some specialist books on the subject. A *word* W in the symbols $a, b, c, \ldots$ is a finite sequence such as $ba^{-}1ccaa^{-1}b^{-1}a$; the empty sequence is denoted by 1. We call two words *equivalent* if one can be obtained from the other by repeatedly replacing xx^{-1} or $x^{-1}x$ by 1 (the empty word) or vice versa. Thus $abb^{-1}a^{-1}c^{-1}$ and $cc^{-1}c^{-1}dd^{-1}$ are both equivalent to c^{-1}. In fact, we shall use the same notation for a word and its equivalence class and so we write simply $abb^{-1}a^{-1}c^{-1} = cc^{-1}c^{-1}dd^{-1} = c^{-1}$. Furthermore, for simplicity $abbc^{-1}c^{-1}c^{-1} = ab^2c^{-3}$, etc. The (equivalence classes of) words form a group if multiplication is defined as juxtaposition: $(ab^{-1}c)(c^{-1}ba) = ab^{-1}cc^{-1}ba = a^2$. Clearly a^{-1} is the inverse of a and $(a^{-1}b^{-1}c)^{-1} = c^{-1}ba$. This group is the *free group* generated by $a, b, c, \ldots$ and it is denoted by $\langle a, b, c, \ldots \rangle$.

Let $R_\mu, R_\nu, \ldots$ be words in the symbols $a, b, c, \ldots$, let $F = \langle a, b, c, \ldots \rangle$ and let K be the normal subgroup of F generated by $R_\mu, R_\nu, \ldots$. Then the quotient group $A = F/K$ is said to be the group generated by $a, b, c, \ldots$ and the *relators* $R_\mu, R_\nu, \ldots$; in notation $A = \langle a, b, c, \ldots | R_\mu, R_\nu, \ldots \rangle$.

Once again we use a word to denote its equivalence class and write equality to express equivalence. More often than not, a group presentation is written with *defining relations* instead of the more pedantic relators. Thus $\langle a, b | a^2 = b^3 \rangle$ denotes the group $\langle a, b | a^2b^{-3} \rangle$. A group is *finitely presented* if in its presentation there are finitely many generators and relations. It is easily seen that two words W_1 and W_2 are equivalent in A iff W_2 can be obtained from W_1 by repeated insertions or deletions of $aa^{-1}, a^{-1}a, bb^{-1}, \ldots$, the relators $R_\mu, R_\nu, \ldots$ and their inverses $R_\mu^{-1}, R_\nu^{-1}, \ldots$. As an example, note that in $A = \langle a, b | a^3b, b^3, a^4 \rangle$ we have $a = aa^3b = a^4b = b$. Hence $1 = a^3b(b^3)^{-1} = a = b$ and so A is the trivial group of order 1.

Even the trivial example above illustrates our difficulties when faced with a group given in terms of defining relations. However, groups defined by their presentations arise naturally in diverse areas of mathematics, especially in knot theory, topology and geometry, so we have to try to overcome the difficulties. The fundamental problems concerning group presentations were formulated by Max Dehn in 1911. These problems ask for general and effective methods for deciding in a finite number of steps (i) whether two given words represent the same group element or not (*word problem*), (ii) whether they represent conjugate elements (*conjugacy problem*) and (iii) whether two finitely presented groups are isomorphic or not (*isomorphism*

problem). All these problems have turned out to be problems in logic, and cannot be solved in general. Explicit solutions of these problems are always based on specific presentations and often make use of group diagrams. (Dehn himself was a particularly enthusiastic advocate of group diagrams.)

Let $A = \langle a, b, \ldots | R_\mu, R_\nu, \ldots \rangle$. We shall attempt to construct the Cayley diagram of A with respect to the generators $a, b, \ldots$. Having got the Cayley diagram, we clearly have a solution to the word problem for this presentation.

The Cayley diagram has the following two properties. (a) The (directed) edges have a *regular* colouring with $a, b, \ldots$, that is for each vertex x and generator g there is exactly one edge coloured g starting at x and exactly one edge coloured g ending at x. (b) Every relation is satisfied at *every vertex*, that is if x is a vertex and R_ν is a relator then the walk starting at x corresponding to R_ν ends at x. How shall we go about finding the Cayley diagram? We try to satisfy (a) and (b) *without ever identifying two vertices unless we are forced to do so*. Thus at each stage we are free to take a *new* vertex and an edge into it (or from it). We identify two vertices when either (b) or (a) force us to do so. If the process stops, we have arrived at the Cayley diagram. Note that until the end we do *not* know that distinct vertices represent distinct group elements.

As an example, let us see how we can find the Cayley diagram of $A = \langle a, b | a^3 = b^2 = (ab)^2 = 1 \rangle$. We replace each double edge corresponding to b by a single undirected edge; this makes the Cayley diagram into a graph. We start with the identity and with a triangle 123 corresponding to $a^3 = 1$; for simplicity we use numbers $1, 2, \ldots$ to denote the vertices, reserving 1 for the identity element. An edge coloured b must start at each of the vertices 1, 2 and 3, giving vertices 4, 5 and 6. Now $a^3 = 1$ must be satisfied at 6, giving another triangle, say 678, whose edges are coloured a, as in Figure VIII.2. At this stage we may care to bring in the relation $(ab)^2 = abab = 1$. Checking it at 8, say, we see that the walk 86314 must end at 8, so the vertices so far denoted by 8 and 4 have to coincide. Next we check $abab = 1$ at 7: the walk $7(8 \equiv 4)/125$ must end at 7 so 5 and 7 are identical. All that remains to check is that the diagram we obtained satisfies (a) and (b), so it is the Cayley diagram of the group in question. In fact, the diagram is exactly the third picture in Figure VIII.1, so the group is S_3.

For $p \geq q \geq r \geq 2$ denote by (p, q, r) the *group* $\langle a, b, c | a^p = b^q = c^r = abc = 1 \rangle$. Given specific values of p, q and r, with a little effort the reader

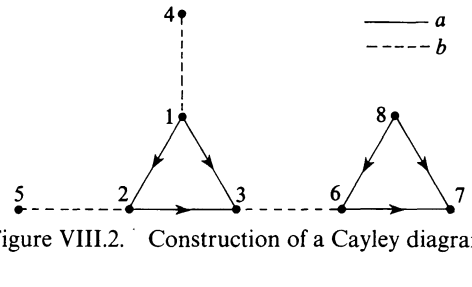


Figure VIII.2. Construction of a Cayley diagram.

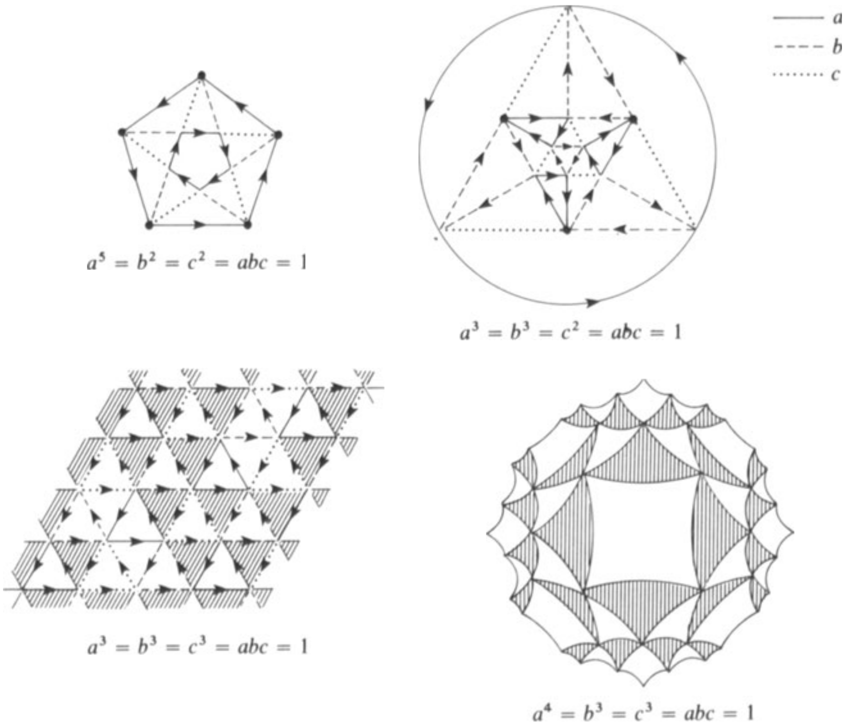


Figure VIII.3. Some Cayley diagrams. The shaded regions correspond to $abc = 1$.

can find the Cayley diagram of the group (p, q, r) with respect to the generators a, b and c. Figure VIII.3 shows some of these diagrams.

The diagrams above indicate some connection with tessellations. The beauty of the use of Cayley diagrams is that one can make good use of this connection. Indeed, the reader who is slightly familiar with tessellations of the sphere, the Euclidean plane and the hyperbolic plane, can easily prove the following result.

Theorem 1. *If* $(1/p) + (1/q) + (1/r) < 1$ *then the group* (p, q, r) *is finite and has order* $2s$ *where* $1/s = (1/p) + (1/q) + (1/r) - 1$. *The Cayley diagram is a tessellation of the sphere* (*as in the first two pictures in Figure* VIII.3).

If $(1/p) + (1/q) + (1/r) \leq 1$ *then the group* (p, q, r) *is infinite. If equality holds, the Cayley diagram is a tessellation of the Euclidean plane, otherwise it is a tessellation of the hyperbolic plane* (*as in the last two pictures in Figure VIII.*3).

As we remarked earlier, groups given by means of their presentations arise frequently in knot theory. In particular, Dehn showed how a presentation of the group of a (tame) *knot* (that is the fundamental group of $\mathbb{R}^3$ after the removal of the knot) can be read off from a projection of the knot

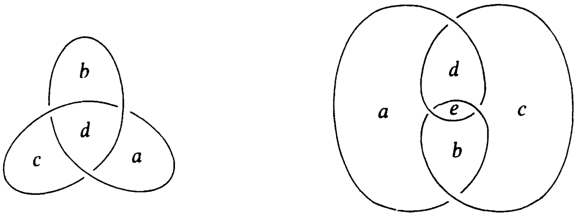


Figure VIII.4. The trefoil and the figure of eight.

into a plane. The projection of the knot forms the boundary of certain bounded domains of the plane. To each of these domains there corresponds a generator (the identity corresponds to the unbounded domain) and to each cross-over there corresponds a relator. The general form of these relators is easily deduced from the two examples shown in Figure VIII.4. (Indeed, readers familiar with the fundamentals of algebraic topology can easily prove the correctness of this presentation.) The group of the *trefoil* (or *clover leaf*) *knot* is $\langle a, b, c, d \mid ad^{-1}b, cd^{-1}a, cbd^{-1}\rangle$ and the group of the *figure of eight knot* is $\langle a, b, c, d, e \mid ab^{-1}c, ad^{-1}eb^{-1}, ed^{-1}cb^{-1}, acd^{-1}\rangle$.

Of course, before embarking on an investigation of the group, it is sensible to attempt to simplify the presentation. For example, $cbd^{-1} = 1$ means that $d = cb$ or $bd^{-1}c = 1$. Thus the group of the trefoil knot is

$$\langle a, b, c, d \mid ad^{-1}b, bd^{-1}c, cd^{-1}a\rangle$$

or, equivalently,

$$\langle a, b, c \mid cb = ba = ac\rangle.$$

We invite the reader to check that the Cayley diagram of this group is made up of replicas of the ladder shown in Figure VIII.5. (Exercise 4). At each edge three ladders are glued together in such a way that when looking at these ladders from above, we see an infinite cubic tree (Figure VIII.5).

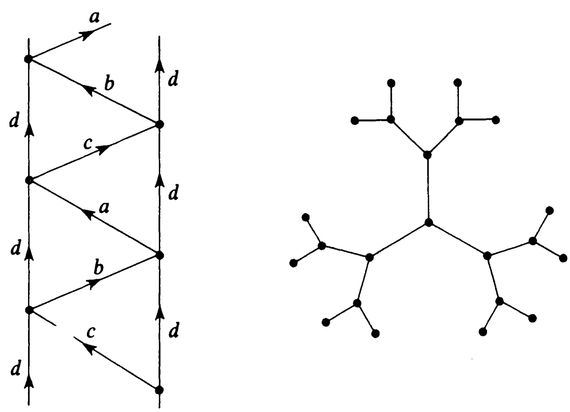


Figure VIII.5. The ingredients of the Cayley diagram of the trefoil knot.

Having obtained the Cayley diagram, we can read off the properties we are interested in. In the case of this group the method does not happen to be too economical, but this is the way Dehn proved in 1910 that the group of the trefoil knot is not the group of a circle, which is infinite cyclic.

Schreier diagrams can be constructed analogously to Cayley diagrams. In fact, in order to determine the structure of a largish group given in a presentation it is often more advantageous to determine first the Schreier diagram of a subgroup. In order to show this, we work through another example, once again due to Dehn. What is the group

$$A = \langle a, b \,|\, a^2 = b^5 = (ba)^3 = 1\rangle?$$

Let us construct the Schreier diagram of the cosets of the subgroup B generated by b. As before, we take a vertex 1 for B and try to construct as big a diagram as conditions (a) and (b) allow us. (Recall that (a) requires the colouring to be regular and (b) requires that *each* defining relation is satisfied at *each* vertex.) However, in this case there is one more condition: the edge coloured b starting at 1 must end in 1 (so it is a loop) since $Bb = B$. Thus after two steps we have the diagram shown in Figure VIII.6. (Once again the edges coloured a will not be directed since $a^2 = 1$.)

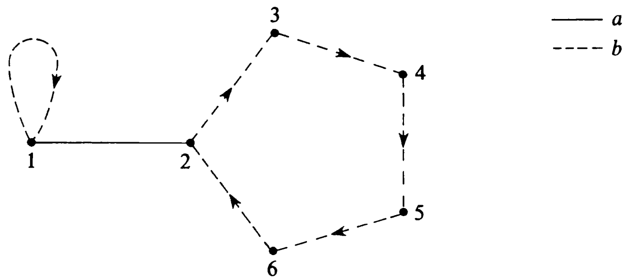


Figure VIII.6. The initial part of the Schreier diagram of A mod B.

Now let us check the condition $bababa = 1$ at vertex 6. The walk $babab$ takes us from 6 to 3, so there must be an edge coloured a from 3 to 6. In order to have edges coloured a starting at 4 and 5, we take up new vertices 7 and 8, together with edges 47 and 58 coloured a. Next we check the condition $ab^{-1}ab^{-1}ab^{-1} = 1$, which is equivalent to $(ba)^3 = 1$ at 7, and find that there is an edge from 7 to 8 coloured b. To satisfy $b^5 = 1$ at 7 we take three new vertices, 9, 10 and 11. Checking $(ba)^3 = 1$ at 11 we find that there is an edge from 9 to 11 coloured a. At this stage we are almost home, but no edge coloured a begins at 10, so we take a new vertex 12 joined to 10 by an edge coloured a. What does the condition $ab^{-1}ab^{-1}ab^{-1} = 1$ tell us at vertex 12? The walk $ab^{-1}ab^{-1}a$ starting at 12 ends at 12, so there must be an edge coloured b *starting* at 12 and *ending* at 12, giving us Figure VIII.7. This is the Schreier diagram we have been looking for, since the colouring is clearly regular and it is a simple matter to check that *each* defining relation

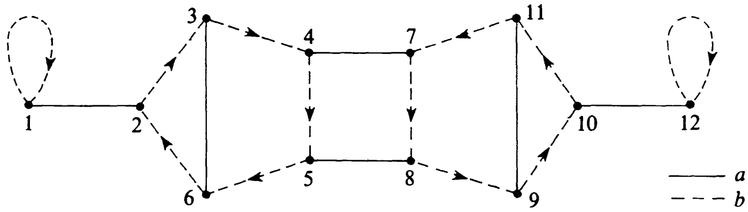


Figure VIII.7. The complete Schreier diagram.

is satisfied at *each* vertex. In fact, $a^2 = 1$ and $b^5 = 1$ are obviously satisfied; since $(ba)^3 = 1$ holds at 6 by construction, it also holds at each vertex of the walk 621236, etc.

This detailed and cumbersome description fails to do justice to the method which, when performed on a piece of paper or on a blackboard, is quick and efficient. The reader is encouraged to find this out for himself.

What the Schreier diagram certainly tells us is the *index* of B: it is simply the number of vertices. Indeed, Schreier diagrams are often constructed on computers just to determine the index of a subgroup. In some cases, as in the example above, it tells us considerably more. The Schreier diagram is essentially a shorthand for the representation of A as a group of permutations of the cosets of B. In this case $a \to (12)(36)(47)(58)(9\ 11)(10\ 12)$ and $b \to (1)(23456)(789\ 10\ 11)(12)$. Since the permutation corresponding to b has order 5, which is exactly the order of b in A, we see that A is a group of order $5.12 = 60$ and is, in fact, isomorphic to A_5, the alternating group on 5 symbols.

If we want our Schreier diagram to carry more information, then we fix certain representatives of the cosets and keep track of the effect of the generators on these representatives. We *decorate* each *coset* by its representative: if H is decorated by $[h]$ then $H = Bh$. Now if there is an edge coloured a from $H = Bh$ to $K = Bk$, then we *decorate* this *edge* by $[\alpha]$ if $ha = \alpha k$. Since $K = Bk = Ha = Bha$, we see that $\alpha \in B$, so the edges are decorated with elements of B. Furthermore, if H, K and L are decorated with h, k and l, and there are edges coloured a, b, c and decorated $[\alpha]$, $[\beta]$ and $[\gamma]$ joining them, as in Figure VIII.8, then $habc = \alpha kbc = \alpha\beta lc = \alpha\beta\gamma h$. In particular,

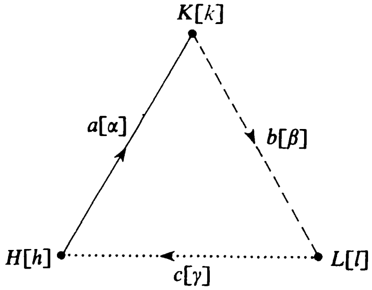


Figure VIII.8. A cycle in a decorated Schreier diagram.

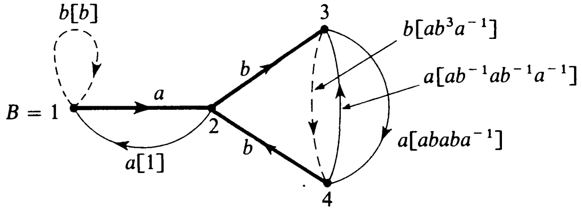


Figure VIII.9. The decorations induced by a spanning tree.

if $h = 1$ we have $abc = \alpha\beta\gamma$. Of course, an analogous assertion holds for arbitrary walks starting and ending at B: the product of the colours equals the product of the decorations.

One of the simplest ways of decorating the vertices and edges makes use of spanning trees. Select a spanning tree of the Schreier diagram. Decorate B, the subgroup itself, by 1 (the identity) and decorate the edges of the spanning tree also by 1. This determines the decoration of *every vertex* (that is coset) and *every edge*. Indeed, for each vertex H the spanning tree contains a unique path from B to H; clearly H has to be decorated with the product $abc\ldots$ corresponding to this path. These coset representatives are said to form a *Schreier system for B mod A*. What is the decoration of a chord HK, an edge not in the tree? By the remark above it is the product of the colours on the B-H path, the edge HK and the K-B path, as in Figure VIII.9.

Since each element of B is the product of the colours on a closed walk from B to B in the Schreier diagram, the decorations of the chords generate B. Thus from a Schreier diagram we can read off a set of generators of B, *independently* of the structure of A.

Theorem 2. *The subgroup B of A is generated by the decorations of the chords.* □

In particular, the subgroup B in Figure VIII.9 is generated by b, ab^3a^{-1}, $ab^{-1}ab^{-1}a^{-1}$ and $ababa^{-1}$.

It is equally simple to find a presentation of B, provided we have a presentation of A. This is obtained by the *Reidemeister–Schreier rewriting process*; we give a quick and loose description of it. The generators of this presentation are the chords of the spanning tree; to distinguish chords of the same colour we write c_i for the edge coloured c starting at vertex i. For each vertex i and each relator R_μ denote by R_μ^i the (word of the) walk starting at i given by R_μ expressed as a product of the c_j, say $R_\mu^i = b_i c_j \ldots$. The reader can easily fill in the missing details in the proof of the following beautiful result, due to Reidemeister and Schreier.

Theorem 3. *The subgroup B has a presentation*

$$\langle a_1, b_1, \ldots, a_2, b_2, \ldots \mid R_\mu^1, R_\nu^1, \ldots, R_\mu^2, R_\nu^2, \ldots \rangle$$ □

Now if we wish to preserve the connection between this presentation of B and the original presentation of A, we simply equate c_i with the decoration of the edge coloured c starting at vertex i.

If A is a free group, its presentation contains no relations. Hence the above presentation of B contains no relations either, so B is also a free group. This is a fundamental result of Nielsen and Schreier.

Theorem 4. *A subgroup of a free group is free. Furthermore, if A is a free group of rank k (that is it has k free generators) and B is a subgroup of index n then B has rank $(k - 1)n + 1$.*

PROOF. The presentation of B given in Theorem 3 is a free presentation on the set of chords of the Schreier diagram. Altogether there are kn edges of which $n - 1$ are tree edges; hence there are $(k - 1)n + 1$ chords.

It is amusing to note that Theorem 3 implies that the rank of a free group is well defined. Indeed, suppose that A is freely generated by $a, b, \ldots$. Then *every* directed multigraph with a distinguished vertex (corresponding to the subgroup) and a regular colouring (by $a, b, \ldots$) is the Schreier diagram of some subgroup B of A, for threre are no relations to be satisfied. Hence subgroups of index n are in 1-1 correspondence with regularly coloured multigraphs of order n. In particular, if A has $k \geq 2$ generators, it has $2^k - 1$ subgroups of index 2, for there are 2^k multigraphs of order 2 regularly coloured with k colours, but one of those is disconnected. The number of subgroups of index 2 is clearly independent of the presentation, so k is determined by A.

§2 Applications of the Adjacency Matrix

Recall from §3, Chapter II that the vertex space $C_0(G)$ of a graph G is the complex vector space of all functions from $V(G)$ into $\mathbb{C}$. Once again we take $V(G) = \{v_1, v_2, \ldots, v_n\}$ so that $\dim C_0(G) = n$ and write the elements of $C_0(G)$ in the form $\mathbf{x} = \sum_{i=1}^n x_i v_i$ or $\mathbf{x} = (x_i)_1^n$; here x_i is the value of $\mathbf{x}$ at v_i, also called the *weight at* v_i. The space $C_0(G)$ is given the natural inner product associated with the basis $(v_i)_1^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i \bar{y}_i$.

We shall concentrate on the adjacency matrix $A = (a_{ij})$ of G, the 0-1 matrix where $a_{ij} = 1$ iff $v_i v_j$ is an edge. As usual, A is identified with a linear endomorphism of $C_0(G)$. To start with, we recollect some simple facts from linear algebra. The matrix A is real and symmetric, so it is *hermitian*, that is $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle$. Hence its *numerical range*

$$V(A) = \{\langle A\mathbf{x}, \mathbf{x} \rangle : \langle \mathbf{x}, \mathbf{x} \rangle = 1\}$$

is a closed interval of the real line. The eigenvalues of A are *real*, say $\lambda_1 > \lambda_2 > \cdots > \lambda_t$, and $V(A)$ is exactly the interval $[\lambda_t, \lambda_1]$. (For simplicity an

eigenvalue of A is said to be an *eigenvalue of* G). We shall write $\lambda_{\max}(G)$ for the maximal eigenvalue λ_1 and $\lambda_{\min}(G)$ for the minimal eigenvalue λ_t. The inner product space $C_0(G)$ has got an orthonormal basis consisting of eigenvectors of A. In particular, if $m(\lambda)$ denotes the (geometric or algebraic) *multiplicity* of an eigenvalue λ then $\sum_{i=1}^{t} m(\lambda_i) = n$. These remarks enable us to deduce some basic properties of the eigenvalues.

Theorem 5. *Let G be a connected graph of order n with adjacency matrix A.*

(i) *Every eigenvalue λ of G satisfies $|\lambda| \le \Delta = \Delta(G)$.*
(ii) *Δ is an eigenvalue of G iff G is regular; if Δ is an eigenvalue then $m(\Delta) = 1$.*
(iii) *If $-\Delta$ is an eigenvalue of G then G is regular and bipartite.*
(iv) *If G is bipartite and λ is an eigenvalue of G then so is $-\lambda$, and $m(\lambda) = m(-\lambda)$.*
(v) $\delta(G) \le \lambda_{\max}(G) \le \Delta(G)$.
(vi) *If H is a spanned subgraph of G then $\lambda_{\min}(G) \le \lambda_{\min}(H) \le \lambda_{\max}(H) \le \lambda_{\max}(G)$.*

PROOF. (i) Let $\mathbf{x} = (x_i)$ be a non-zero eigenvector with eigenvalue λ. Let x_p be a weight with maximum modulus: $|x_p| \ge |x_i|$ for every i; we may assume without loss of generality that $x_p = 1$. Then

$$|\lambda| = |\lambda x_p| = \left|\sum_{l=1}^{n} a_{pl} x_l\right| \le \sum_{l=1}^{n} a_{pl}|x_l| \le |x_p| d(v_p) \le |x_p|\Delta = \Delta, \quad (*)$$

showing $|\lambda| \le \Delta$.

(ii) If $\lambda = \Delta$ is an eigenvalue and $\mathbf{x}$, x_p are chosen as in (i), then (*) implies that $d(v_p) = \Delta$ and $x_l = x_p = 1$ whenever v_l is adjacent to v_p. In turn this implies that $d(v_l) = \Delta$ and $x_k = x_l = 1$ whenever v_k is adjacent to v_l, and so on; as G is connected $d(v_i) = \Delta$ and $x_i = 1$ for every i. Hence G is Δ-regular and $\mathbf{x}$ is $\mathbf{j}$, the vector all of whose entries are 1.

Conversely, if G is Δ-regular then $(A\mathbf{j})_i = \sum_{l=1}^{n} a_{il} = \Delta$ so $A\mathbf{j} = \Delta\mathbf{j}$.

(iii) If $\lambda = -\Delta$ is an eigenvalue then, as in (ii) inequality (*) implies that $d(v_p) = \Delta$ and $x_l = -x_p = -1$ whenever v_l is adjacent to v_p. As in (ii), this implies that G is Δ-regular. Furthermore, at each vertex v_k adjacent to v_l the weight is 1, at each neighbour of v_k it is -1, and so on. The weight is 1 at the vertices at an even distance from v_p and it is -1 at the other vertices and every edge joins vertices of different weights. Thus G is bipartite, say $V = V_1 \cup V_2$, where $x_p \in V_1$.

(iv) Suppose G is bipartite with vertex classes V_1 and V_2. Let $\mathbf{b}$ be the function (vector) that is 1 on V_1 and -1 on V_2. Then $\mathbf{x} \to \mathbf{bx} = (b_i x_i)_1^n$ is an automorphism of the vector space $\mathbb{C}_0(G)$. Now if $A\mathbf{x} = \lambda\mathbf{x}$ and $v_i \in V_1$, say, then

$$(A(\mathbf{bx}))_i = \sum_{j=1}^{n} a_{ij} b_j x_j = \sum_{v_j \in V_1} a_{ij} x_j - \sum_{v_j \in V_2} a_{ij} x_j = -\sum_{v_j \in V_2} a_{ij} x_j$$

$$= -\sum_{v_j \in V_1} a_{ij} x_j - \sum_{v_j \in V_2} a_{ij} x_j = -\sum_{j=1}^{n} a_{ij} x_j = -\lambda x_i = -\lambda(\mathbf{bx})_i.$$

Hence **b** gives an isomorphism between $\ker(A - \lambda)$ and $\ker(A + \lambda)$, showing $m(\lambda) = m(-\lambda)$.

(v) We know already that $\lambda_{\max}(G) \le \Delta(G)$. Note that for $\mathbf{j} = (1, 1, \ldots, 1)$ we have $\langle \mathbf{j}, \mathbf{j} \rangle = n$ so

$$V(A) \ni \frac{1}{n} \langle A\mathbf{j}, \mathbf{j} \rangle = \frac{1}{n} \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl} = \frac{1}{n} \sum_{k=1}^{n} d(v_k) \ge \delta(G).$$

Hence $\lambda_{\max}(G) = \max V(A) \ge \delta(G)$.

(vi) It suffices to prove the result for a spanned subgraph H of order $n - 1$, say $V(H) = \{v_1, v_2, \ldots, v_{n-1}\}$.

Let A' be the adjacency matrix of H. Then there is a vector $\mathbf{y} \in \mathbb{C}_0(H)$ such that $\langle \mathbf{y}, \mathbf{y} \rangle = 1$ and $\langle A'\mathbf{y}, \mathbf{y} \rangle = \lambda_{\max}(H)$. Let $\mathbf{x} = (y_1, y_2, \ldots, y_{n-1}, 0)$. Then $\mathbf{x} \in \mathbb{C}_0(G)$, $\langle \mathbf{x}, \mathbf{x} \rangle = 1$ and $\langle A\mathbf{x}, \mathbf{x} \rangle = \langle A'\mathbf{y}, \mathbf{y} \rangle = \lambda_{\max}(H) \in V(A)$. Consequently $\lambda_{\max}(G) \ge \lambda_{\max}(H)$. The other inequality is proved analogously. □

It is reasonable to expect that a graph with many automorphisms will have particularly pleasing properties and that these will be reflected in the adjacency matrix. The *automorphism group of a graph* G is the group, Aut G, of permutations of the vertices preserving adjacency. Every abstract group can be represented as the automorphism group of some graph. For instance, if F is any finite group, consider its Cayley diagram with respect to some set of generators. The automorphism group of this *coloured* and *directed multigraph* is exactly F. It only remains to replace each edge of this diagram by a suitable subgraph which bears the information previously given by the direction and colour. This produces a graph G with automorphism group isomorphic to F. An example is shown in Figure VIII.10.

Each $\pi \in \text{Aut } G$ induces an endomorphism of $C_0(G)$ which is described by a permutation matrix P. In fact, an arbitrary permutation matrix Q corresponds to an automorphism of G precisely when it commutes with the

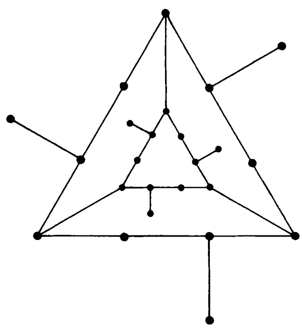


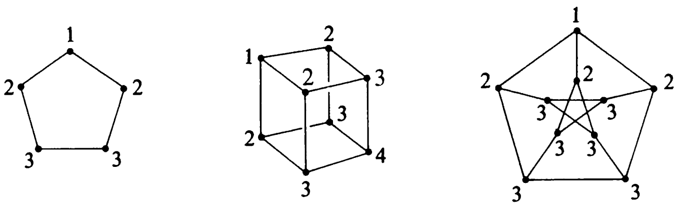
Figure VIII.10. A graph with automorphism group S_3, constructed from the Cayley diagram in Figure VIII.1.

adjacency matrix A, that is $AQ = QA$. The group of these matrices therefore faithfully represents Aut G. Regarding A as an endomorphism of $C_0(G)$, we find that the *eigenspaces* of A are *invariant* under P. In particular, if an eigenvalue of A is simple (that is it has multiplicity 1) then P must map an eigenvector to a multiple of itself. Thus if all eigenvalues are simple, we must have $P^2 = I$. This is a strong restriction on those permutations which might correspond to automorphisms of G. For example, if G has at least 3 vertices and every eigenvalue is simple, then Aut G cannot be vertex-transitive, so not every pair of vertices can be interchanged by an automorphism.

These remarks indicate how the methods of representation theory may be used to deduce restrictions on the adjacency matrix of graphs which have extensive automorphism groups. The lack of space prevents us from exploring this further. Instead, we shall use algebraic methods to study graphs which are highly regular although this regularity is not expressed in terms of the automorphism group.

Algebraic methods are particularly useful if we want to prove that certain regularity conditions cannot be satisfied except perhaps for a small set of parameters. A problem of this type arose in Chapter IV: for which values of k is there a k-regular graph of order $n = k^2 + 1$ and girth 5? We shall show later that if there is such a graph then k is 2, 3, 7 or 57. Group theory is particularly rich in problems of this type, especially because of the search for sporadic simple groups. Later on we shall mention some examples.

Regularity, that is the condition that all vertices have the same degree k, is not sufficiently restrictive. There are too many k-regular graphs of order n (provided $k < n$ and kn is even) so we have to go considerably further to find interesting classes. Call a connected graph G *highly regular with collapsed adjacency matrix* $C = (c_{ij})$ if for every vertex $x \in V = V(G)$ there is a partition of V into sets $V_1 = \{x\}, V_2, \ldots, V_p$ such that each vertex $y \in V_j$ is adjacent to exactly c_{ij} vertices in V_i (see Figure VIII.11). It is immediate from the definition that G is regular, say every vertex has degree k. In this case each *column sum* in the collapsed matrix is k. The collapsed matrix C



$$\begin{pmatrix} 0 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 3 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

Figure VIII.11. The pentagon, the cube and the Petersen graph together with collapsed adjacency matrices.

can be obtained from the adjacency matrix A as follows:

$$c_{ij} = \sum_{v_s \in V_i} a_{st}, \quad \text{where } v_t \in V_j.$$

The point is exactly that the above sum is *independent* of the representative v_t of V_i. We are especially interested in the collapsed adjacency matrix C if it is of a much smaller size than A.

Consider a partition $V_1 = \{v_1\}, V_2, \ldots, V_p$ belonging to the vertex $x = v_1$. Let P be the p-dimensional complex vector space of formal sums $\sum_{i=1}^{p} z_i V_i$, $z_i \in \mathbb{C}$, and let $\pi\colon \mathbb{C}_0(G) \to P$ be the linear map given by

$$\pi\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{i=1}^{n} \left(\sum_{v_j \in V_i} x_j\right) V_i.$$

Of course, C can be considered to map P into itself.

Theorem 6.

(i) $\pi A = C\pi$, *that is* $\pi(A\mathbf{x}) = C(\pi\mathbf{x})$ *for every* $\mathbf{x} \in \mathbb{C}_0(G)$.

(ii) *The adjacency matrix A and the collapsed adjacency matrix C have the same minimal polynomial. In particular, λ is an eigenvalue of A iff it is a root of the characteristic polynomial of C.*

PROOF. (i) Let us show that $\pi(Av_t) = C(\pi v_t)$, where v_t is the basis vector corresponding to an arbitrary vertex $v_t \in V_j$. To do this it suffices to check that the ith coordinates of the two sides are equal. Clearly $\pi v_t = V_j$ so

$$(\pi(Av_t))_i = \sum_{v_s \in V_i} a_{st}$$

and

$$(C(\pi v_t))_i = (CV_j)_i$$

and these are equal by definition.

(ii) Let q be the minimal polynomial of C. In order to prove $q(A) = 0$, by symmetry it suffices to check that $(q(A)\mathbf{x})_1 = 0$ for every $\mathbf{x}$. This is indeed so:

$$(q(A)\mathbf{x})_1 = (\pi(q(A)\mathbf{x}))_1 = (q(C)(\pi\mathbf{x}))_1 = (\mathbf{0})_1 = 0.$$

Conversely, the minimal polynomial of A annihilates C since $\pi\mathbb{C}_0(G) = P$. □

This result enables us to restrict rather severely the matrices C that may arise as collapsed adjacency matrices.

Theorem 7. *Let G be a connected non-bipartite highly regular graph of order n with collapsed adjacency matrix C. Let $\lambda_1, \lambda_2, \ldots, \lambda_r$ be the roots of the*

characteristic polynomial of C different from k, the degree of the vertices of G. Then there are natural numbers $m_1, m_2, \ldots, m_r$ such that

$$\sum_{i=1}^{r} m_i = n - 1$$

and

$$\sum_{i=1}^{r} m_i \lambda_i = -k.$$

PROOF. We know from Theorem 5 that $\lambda_1, \lambda_2, \ldots, \lambda_r$ are the eigenvalues of A in addition to k, which has multiplicity 1. Thus if $m(\lambda_i)$ is the multiplicity of λ_i then

$$1 + \sum_{i=1}^{r} m(\lambda_i) = n,$$

since $C_0(G)$ has an orthonormal basis consisting of eigenvectors of A. Furthermore, since the trace of A is 0 and a change of basis does not alter the trace,

$$\operatorname{tr} A = k + \sum_{i=1}^{r} m(\lambda_i)\lambda_i = 0. \qquad \square$$

The condition expressed in Theorem 7 is not easily satisfied, especially if $\lambda_1, \lambda_2, \ldots, \lambda_r$ are not rational numbers, so it does rule out the possibility of constructing many highly regular graphs with seemingly feasible parameters. For the so-called strongly regular graphs we shall rewrite the condition in a more attractive form. A graph G is said to be *strongly regular with parameters* (k, a, b) if it is k-regular, any two adjacent vertices have exactly a common neighbours and any two non-adjacent vertices have $b \geq 1$ common neighbours. In other words, G is highly regular with collapsed adjacency matrix

$$C = \begin{pmatrix} 0 & 1 & 0 \\ k & a & b \\ 0 & k-a-1 & k-b \end{pmatrix}.$$

Theorem 8. *If there is a strongly regular graph of order n with parameters (k, a, b) then*

$$m_1, m_2 = \frac{1}{2}\left\{n - 1 \pm \frac{(n-1)(b-a) - 2k}{\{(a-b)^2 + 4(k-b)\}^{1/2}}\right\}$$

are natural numbers.

PROOF. The characteristic polynomial of the collapsed adjacency matrix C is

$$x^3 + (b - a - k)x^2 + ((a - b)k + b - k)x + k(k - b).$$

On dividing by $x - k$, we find that the roots different from k are

$$\lambda_1, \lambda_2 = \tfrac{1}{2}\{a - b \pm \{(a - b)^2 + 4(k - b)\}^{1/2}\}.$$

By theorem 7 there are natural numbers m_1 and m_2 satisfying

$$m_1 + m_2 = n - 1$$

and

$$m_1\lambda_1 + m_2\lambda_2 = -k.$$

Solving these for m_1 and m_2 we arrive at the assertion of the theorem. □

Theorem 8 is sometimes called the *rationality condition* for strongly regular graphs. It is also easily proved without invoking Theorem 7. Indeed, if A is the adjacency matrix of a strongly regular graph with parameter (n, k, a, b), then

$$(A^2)_{ij} = \begin{cases} k & \text{if } i = j, \\ a & \text{if } v_iv_j \in E(G), \\ b & \text{otherwise} \end{cases}$$

Hence

$$A^2 = kI + aA + b(J - I - A),$$

where J is the matrix all whose entries are 1. Therefore J is a quadratic polynomial in A, so J and A are simultaneously diagonalizable. Noting that J has only two eigenvalues: n, with multiplicity 1, and 0, with multiplicity $n - 1$, one can easily find that λ_1 and λ_2 are as above (cf. Exercise 28).

Now we shall use the rationality condition to fulfil our promise concerning Moore graphs of diameter 2 (or girth 5).

Theorem 9. *Suppose there is a k-regular graph G of order $n = k^2 + 1$ and diameter 2. Then $k = 2, 3, 7$ or 57.*

Proof. We know from Theorem 1 of Chapter IV that G is strongly regular with parameters $(k, 0, 1)$. By the rationality condition at least one of the following two conditions has to hold:

(i) $(n - 1) - 2k = k^2 - 2k = 0$ and $n - 1 = k^2$ is even,
(ii) $1 + 4(k - 1) = 4k - 3$ is a square, say $4k - 3 = s^2$.

Now if (i) holds then $k = 2$.

If (ii) holds then $k = \tfrac{1}{4}(s^2 + 3)$; on putting this into the expression for the multiplicity m_1 we find that

$$m_1 = \frac{1}{2}\left\{\frac{1}{16}(s^2 + 3)^2 + \frac{[(s^2 + 3)^2/16] - [(s^2 + 3)/2]}{s}\right\},$$

that is

$$s^5 + s^4 + 6s^3 - 2s^2 + (9 - 32m_1)s - 15 = 0.$$

Hence s divides 15 so s is one of the values 1, 3, 5 and 15, giving $k = 1, 3, 7$ or 57. The case $k = 1$ is clearly unrealizable. □

It is worth noting that for $k = 2$, 3 and 7 there are unique k-regular graphs of order $k^2 + 1$ and diameter 2 (and so girth 5). In particular for $k = 2$ it is a pentagon and for $k = 3$ it is the Petersen graph. However, it is not known whether or not $k = 57$ can be realized.

Sporadic simple groups are those simple groups which do not belong to one of the infinite sequences consisting of cyclic groups of prime order, alternating groups of degree at least 5 and simple groups of Lie type. As we remarked earlier, sporadic simple groups are often connected to strongly regular graphs. For example, there is a strongly regular graph with parameters (162, 105, 81) and the Maclaughlin group of order 898 128 000 is a subgroup of index 2 of the automorphism group of this graph. Similarly, there is a strongly regular graph with parameters (416, 100, 96) and the Suzuki group, which is a simple group of order 448 345 497 600, is a subgroup of index 2 of the automorphism group of this graph.

§3 Enumeration and Pólya's Theorem

We cannot end the book without considering perhaps the most basic question about graphs, namely, how many of them are there? We may want to count graphs with a given set of vertices or we may be interested in the number of isomorphism classes of certain graphs. As we saw in Chapter VII, counting labelled graphs is relatively easy, for instance there are $2^{\binom{n}{2}} = 2^N$ labelled graphs on n vertices, of which $\binom{N}{m}$ have m edges. Furthermore, by applying Theorem 8 of Chapter II to the complete graph, one can easily show that there are n^{n-2} labelled trees of order n. This result was first obtained by Cayley; we present it here with a proof due to Prüfer, which is independent of Theorem 8 of Chapter II.

Theorem 10. *There are n^{n-2} trees on n labelled vertices.*

Proof. As in Chapter VII, let $V = \{1, 2, \ldots, n\}$ be the set of vertices. Given a tree T, associate a code with T as follows. Remove the endvertex with the smallest label and write down the label of the adjacent vertex. Repeat the process until only two vertices remain. The code obtained is a sequence of length $n - 2$ consisting of some numbers from $1, 2, \ldots, n$; of course any number may occur several times in the code (see Figure VIII.12). As the reader should check, each of the n^{n-2} possible codes corresponds to a unique tree. □

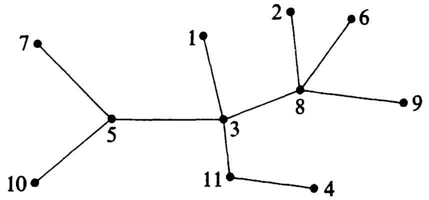


Figure VIII.12. The Prüfer code of this tree is (3, 8, 11, 8, 5, 8, 3, 5, 3).

It is easily seen that the label of a vertex of degree d occurs exactly $d - 1$ times in the Prüfer code of the tree. Thus the proof has the following consequence.

Corollary 11. *Let $d_1 \leq d_2 \leq \cdots \leq d_n$ be the degree sequence of a tree: $d_1 \geq 1$ and $\sum_{i=1}^n d_i = 2n - 2$. Then the number of labelled trees of order n with degree sequence $(d_i)_1^n$ is given by the multinomial coefficient*

$$\binom{n-2}{d_1 - 1, d_2 - 1, \ldots, d_n - 1}. \qquad \square$$

The difficulties we encounter change entirely if we wish to count certain graphs within isomorphism. Given graphs G_1 and G_2 with a common vertex set V, when are they isomorphic? They are isomorphic if there is a permutation π of V which maps G_1 onto G_2. Of course, strictly speaking π does not act on graphs, it only *induces* a permutation α of $X = V^{(2)}$, the *pairs* of vertices, and it is α that maps an *edge* of G_1 into an *edge* of G_2 and a *non-edge* of G_1 into a *non-edge* of G_2. Now G_i is naturally identified with a subset of X or, equivalently, with a function $f_i: X \to \{0, 1\}$. Therefore G_1 is isomorphic to G_2 iff there is a permutation α of $X = V^{(2)}$ (coming from a permutation π of V) such that $\alpha^* f_1 = f_2$, where α^* is the permutation of the set of functions $\{0, 1\}^X$ induced by α. Thus counting graphs within isomorphism is a special case of the following problem. Given sets X and Y, and a group Γ acting on X, let Γ act on the set of functions Y^X in the natural way. How many orbits are there in Y^X? The main aim of this section is to present a beautiful theorem of Pólya that answers this question.

Let Γ be a group of permutations acting on a (finite) set X. For $x, y \in X$ put $x \sim y$ iff $y = \alpha x$ for some $\alpha \in \Gamma$. Then $\sim$ is an equivalence relation on X; if $x \sim y$ we say that x is *equivalent to* y *under* Γ. The equivalence class of x is called the Γ-*orbit* (or simply *orbit*) of x, and is denoted by $[x]$. For $x, y \in X$ put

$$\Gamma(x, y) = \{\alpha \in \Gamma: \alpha x = y\}.$$

Of course, $\Gamma(x, y)$ is non-empty iff $[x] = [y]$, that is x and y belong to the same orbit. $\Gamma(x) = \Gamma(x, x)$ is the *stabiliser* of x; it is a subgroup of Γ. Note that if $y = \beta x$ then

$$\Gamma(x, y) = \{\alpha: \alpha x = y\} = \{\alpha: \alpha x = \beta x\} = \{\alpha: \beta^{-1}\alpha \in \Gamma(x)\} = \beta\Gamma(x),$$

so $\Gamma(x, y)$ is a coset of $\Gamma(x)$. We see that $|\Gamma(x)|$ depends on the equivalence class of x, so we may put $s([x]) = |\Gamma(x)|$. Clearly

$$\Gamma = \bigcup_{y \in [x]} \Gamma(x, y),$$

and this gives us

$$|\Gamma| = |[x]||\Gamma(x)| = |[x]|s([x]).$$

The theorem mentioned in the title of this section is based on a version of a lemma due to Burnside concerning the sum of the "weights" of orbits. Let $O_1, \ldots, O_l$ be the Γ-orbits, let A be an arbitrary Abelian group (written additively) and let $w: X \to A$ be a function that is constant on orbits. We call w a *weight function* and define the weight of O_i by $w(O_i) = w(x)$, $x \in O_i$. For a permutation $\alpha \in \Gamma$ we denote by $F(\alpha)$ the set of elements fixed by α, that is $F(\alpha) = \{x \in X : \alpha x = x\}$. Thus $x \in F(\alpha)$ iff $\alpha \in \Gamma(x)$.

Lemma 12. $|\Gamma| \sum_{i=1}^{l} w(O_i) = \sum_{\alpha \in \Gamma} \sum_{x \in F(\alpha)} w(x)$.

PROOF.

$$\sum_{\alpha \in \Gamma} \sum_{x \in F(\alpha)} w(x) = \sum_{x \in X} \sum_{\alpha \in \Gamma(x)} w(x) = \sum_{i=1}^{l} \sum_{x \in O_i} \sum_{\alpha \in \Gamma(x)} w(x)$$

$$= \sum_{i=1}^{l} w(O_i) \sum_{x \in O_i} \sum_{\alpha \in \Gamma(x)} 1 = \sum_{i=1}^{l} w(O_i)|O_i|s(O_i) = |\Gamma| \sum_{i=1}^{l} w(O_i).$$

□

The original form of Burnside's lemma is obtained on choosing $A = \mathbb{Z}$ and $w \equiv 1$:

$$N(\Gamma) = \frac{1}{|\Gamma|} \sum_{\alpha \in \Gamma} \sum_{x \in F(\alpha)} 1 = \frac{1}{|\Gamma|} \sum_{\alpha \in \Gamma} |F(\alpha)|,$$

where $N(\Gamma) = l$ is the *number of orbits*.

We shall illustrate by two very simple examples, that even Burnside's lemma can be used to calculate the number of equivalence classes of certain objects.

EXAMPLE 1. Let $X = \{1, 2, 3, 4\}$ and $\Gamma = \{1, (12), (34), (12)(34)\}$. What is $N(\Gamma)$? Clearly $F(1) = \{1, 2, 3, 4\}$, $F((12)) = \{3, 4\}$, $F((34)) = \{1, 2\}$ and $F((12)(34)) = \varnothing$. Thus $N(\Gamma) = \frac{1}{4}\{4 + 2 + 2 + 0\} = 2$.

EXAMPLE 2. Consider all bracelets made up of 5 beads. The beads can be red, blue and green, and two bracelets are considered to be identical if one can be obtained from the other by rotation. (Reflections are not allowed!) How many distinct bracelets are there?

In this case we choose X to be the set of all $3^5 = 243$ bracelets and let Γ be C_5, the cyclic group of order 5, acting on X. Then the question is: how many orbits does Γ have? For the identity $1 \in \Gamma$ clearly $F(1) = X$. For every non-trivial rotation $\alpha \in \Gamma$ only the 3 monochromatic bracelets are invariant under α, so $N(\Gamma) = \frac{1}{5}\{243 + 3 + 3 + 3 + 3\} = 51$.

The second example resembles a little the problem we really want to tackle. Let Γ be a group of permutations of a (finite) set D. Let R be another (finite) set and let us consider the set R^D *of all functions from D into R*. Each $\alpha \in \Gamma$ can be made to act on R^D; namely define $\alpha^*: R^D \to R^D$ by

$$(\alpha^* f)(d) = f(\alpha d), \qquad f \in R^D, d \in D.$$

Then

$$\Gamma^* = \{\alpha^*: \alpha \in \Gamma\}$$

is a group of permutations of R^D; as an abstract group, it is isomorphic to Γ and we distinguish it from Γ only to emphasize that it acts on R^D while Γ acts on D.

As customary in connection with Pólya's theorem, we adopt an intuitive terminology. D is called the *domain* and its elements are *places*, R is the *range* and its elements are *figures*, the functions in R^D are called *configurations*, finally a *pattern* is an equivalence class of configurations under Γ^*, that is a Γ^*-orbit. Our main aim is to calculate the number of distinct patterns.

The origin of this terminology is that a function $f \in R^D$ is an arrangement of some figures into the places in such a way that for each place there is exactly one figure in that place, but each figure can be put into as many places as we like. Two configurations mapped into each other by an element of Γ^* have the same pattern and are not distinguished. Thus in the second example the places are, say, 1, 2, 3, 4 and 5, the figures are r, b and g (for red, blue and green) and a configuration is a sequence of the type g, b, b, r, b, that is a bracelet. The group Γ is generated by (12345) and distinct patterns correspond to distinguishable bracelets.

In addition to counting the number of distinct patterns, we may wish to count the number of patterns of a certain type. It turns out that all these problems can be solved at once, provided we learn enough about the cycle structure of permutations in Γ acting on D, and are willing to store much information about the patterns.

Each element $\alpha \in \Gamma$ is an essentially unique product of disjoint cycles (cyclic permutations) acting on D. If $\alpha = \xi_1 \xi_2 \dots \xi_m$ is such a product, we say that $\xi_1, \dots, \xi_m$ *are the cycles of* α. In the product we include cycles of length 1 as well so that every $\alpha \in D$ appears in exactly one cycle; if $|\xi|$ denotes the number of elements in ξ then $\sum_{k=1}^{m} |\xi_k| = d$, where $d = |D|$ is the number of elements in D. Denote by $j_k(\alpha)$ the number of cycles of α having length k; by the previous equality $\sum_{k=1}^{m} k j_k(\alpha) = d$. Note that $|F(\alpha)|$, appearing in Burnside's lemma, is exactly $j_1(\alpha)$, the number of elements of D fixed by α.

We define the *cycle sum* of Γ to be

$$\tilde{Z}(\Gamma; a_1, \ldots, a_d) = \sum_{\alpha \in \Gamma} \prod_{k=1}^{d} a_k^{j_k(\alpha)},$$

where, as before $d = |D|$. The reader should bear in mind that $\tilde{Z}$ depends on the *action of* Γ *on* D, not only on the abstract group Γ. Note also that the cycle sum is a *polynomial* in $a_1, a_2, \ldots, a_d$ with integer coefficients; it tells us the distribution of cycles in the elements of Γ. When writing down a cycle sum, it is useful to remember that $\sum_{k=1}^{d} kj_k(\alpha) = d$ for every α. The customary *cycle index* of Γ is $Z(\Gamma; a_1, \ldots, a_d) = (1/|\Gamma|)\tilde{Z}(\Gamma; a_1, \ldots, a_d)$. As we shall consider general rings instead of the more usual polynomial ring with rational coefficients, we have to use the cycle sum since we cannot divide by $|\Gamma|$.

Let A be an arbitrary *commutative ring* and let $w: R \to A$ be a function. We call $w(r)$ the *weight* of the figure r, and for $k = 1, 2, \ldots$ define the kth *figure sum* as

$$s_k = \sum_{r \in R} w(r)^k.$$

Furthermore, the *weight* of a configuration $f \in R^D$ is

$$w(f) = \prod_{a \in D} w(f(a)).$$

Clearly any two configurations equivalent under Γ^* have the same weight, so we may define the weight of a pattern O_i by

$$w(O_i) = w(f), \qquad f \in O_i.$$

Our aim is to learn about the *pattern sum*

$$S = \sum_{i=1}^{l} w(O_i),$$

where $O_1, O_2, \ldots, O_l$ are the Γ^*-orbits, that is the distinct patterns.

Note that $w(r)$, $w(f)$, s_k and S are all elements of our commutative ring A. If we have a way of determining the pattern sum S, it is up to us to choose A and the weight function $w: R \to A$ in such a way that S can be "decoded" to tell us all we want to know about various sets of patterns. In practice one always chooses A to be a polynomial ring ($\mathbb{Z}[x]$, $Q[x, y]$, etc.), and usually $w(r)$ is a monic polynomial; the information we look for is then given by certain coefficients of the polynomial S. We shall give several examples after the proof of our main result, Pólya's enumeration theorem.

Theorem 13. (Pólya's enumeration theorem)

$$|\Gamma| S = \tilde{Z}(\Gamma; s_1, s_2, \ldots, s_d).$$

PROOF. By Lemma 12

$$|\Gamma|S = |\Gamma|\sum_{i=1}^{l} w(O_i) = \sum_{\alpha\in\Gamma}\ \sum_{f\in F(\alpha^*)} w(f).$$

Now clearly

$F(\alpha^*) = \{f \in R^D : f \text{ is constant on cycles of } \alpha\}$, so if $\xi_1, \xi_2, \ldots, \xi_m$ are the cycles of α and $a \in \xi_i$ means that a is an element of the cycle ξ_i then

$$F(\alpha^*) = \{f \in R^D : r_i \in R \quad \text{and} \quad f(a) = r_i \quad \text{if} \quad a \in \xi_i,\ i = 1, 2, \ldots, m\}.$$

Hence

$$\sum_{f\in F(\alpha^*)} w(f) = \sum_{(r_i)\subset R}\ \prod_{i=1}^{m} w(r_i)^{|\xi_i|} = \prod_{k=1}^{d}\left(\sum_{r\in R} w(r)^k\right)^{j_k(\alpha)} = \prod_{k=1}^{d} s_k^{j_k(\alpha)},$$

giving

$$|\Gamma|S = \sum_{\alpha\in\Gamma}\ \prod_{k=1}^{d} s_k^{j_k(\alpha)} = \tilde{Z}(\Gamma; s_1, s_2, \ldots, s_d). \qquad \square$$

If $|\Gamma|$ has an inverse in the ring A, say if A is a polynomial ring over the rationals, then Theorem 13 can also be written in its more usual form:

$$S = Z(\Gamma; s_1, s_2, \ldots, s_d).$$

Let us illustrate now how the theorem can be applied.

EXAMPLE 3. Let us consider again the bracelets made up of five beads, which can be red, blue and green. Then $D = \{1, 2, 3, 4, 5\}$ is the set of places of the beads, $R = \{r, b, g\}$ is the set of colours (figures) and Γ is C_5, the cyclic group of order 5 generated by the permutation (12345). The cycle sum is $\tilde{Z} = a_1^5 + 4a_5$. On choosing $A = \mathbb{Z}$ and $w(r) = w(b) = w(g) = 1$, we find that $s_k = 3$ for every k, so $5S = 3^5 + 4.3$. Since each pattern (bracelet) has weight 1, there are $\frac{1}{5}\{3^5 + 12\} = 51$ distinct patterns (bracelets).

On choosing $A = \mathbb{Z}[x, y]$ and $w(r) = 1$, $w(b) = x$, $w(g) = y$, we find that $S = \frac{1}{5}\{(1 + x + y)^5 + 4(1 + x^5 + y^5)\}$. Now it is easy to extract much information from this form of S. For example, a bracelet has weight xy^2 iff it has 2 red, 1 blue and 2 green beads. Thus the number of such bracelets is the coefficient of xy^2 in the polynomial S, that is $(1/5)(5!/2!2!) = 6$.

EXAMPLE 4. What happens if in the previous example we allow reflections? Then Γ is the dihedral group D_5, the group of symmetries of the regular pentagon, whose cycle sum is $a_1^5 + 4a_5 + 5a_1a_2^2$. Thus if we take, as before, $A = \mathbb{Z}[x, y]$, $w(r) = 1$, $w(b) = x$ and $w(g) = y$, we find that the number of bracelets containing 2 red, 1 blue and 2 green beads is the coefficient of xy^2 in $\frac{1}{10}\{(1 + x + y)^5 + 4(1 + x^5 + y^5) + 5(1 + x + y)(1 + x^2 + y^2)^2\}$, that is, $3 + 1 = 4$.

EXAMPLE 5. This is the example Pólya used to illustrate his theorem. Place 3 red, 2 blue and 1 yellow balls in the 6 vertices of an octahedron. In how many distinct ways can this be done? The group of symmetries of the octahedron has cycle sum $a_1^6 + 6a_1^2 a_4 + 3a_1^2 a_2^2 + 6a_2^3 + 8a_3^2$: a_1^6 comes from the identity, $6a_1^2 a_4$ from rotations through π about axes through the vertices, $6a_2^3$ from rotations through π about axes through midpoints of edges, and, finally, a_3^2 is the summand corresponding to a rotation through $2\pi/3$ about an axis going through the centre of a face. On taking $A = \mathbb{Z}[x, y]$, $w(r) = 1$, $w(b) = x$ and $w(y) = y$, we see that the required number is the coefficient of $x^2 y$ in

$$\tfrac{1}{24}\{(1 + x + y)^6 + 6(1 + x + y)^2(1 + x^4 + y^4)$$
$$+ 3(1 + x + y)^2(1 + x^2 + y^2)^2 + 6(1 + x^2 + y^2)^3 + 8(1 + x^3 + y^3)^2\},$$

that is, 3.

It should be clear by now that the theorem loses nothing from its generality if instead of a general commutative ring A we take $\mathbb{Z}[x_r: r \in R]$, the polynomial ring over the integers in variables with the elements of R, and we define the weight function as $w(r) = x_r$. Then the pattern sum S contains all the information the theorem can ever give us. In particular, if $w: R \to A$ is an arbitrary weight function then the corresponding pattern sum is obtained by replacing x_r by $w(r)$ in S. However, if R is large, the calculations may get out of hand if we do not choose a "smaller" ring than $\mathbb{Z}[x_r: r \in R]$, which is tailor-made for the problem at hand. The choice of a smaller ring is, of course, equivalent to a substitution into S.

EXAMPLE 6. Place red, blue, green and yellow balls into the vertices of an octahedron. Denote by P_i the set of patterns in which the total number of red and blue balls is congruent to i modulo 4. What is $|P_0| - |P_2|$?

The cycle sum of the rotation group of the octahedron was calculated in Exercise 5 and was found to be $a_1^6 + 6a_1^2 a_4 + 3a_1^2 a_2^2 + 6a_2^3 + 8a_3^2$.

Let $A = \mathbb{C}$, the field of complex numbers, and put $w(r) = w(b) = i$, $w(g) = w(y) = 1$. Then for a pattern f we have $\operatorname{Re} w(f) = 1$ if $f \in P_0$, $\operatorname{Re} w(f) = -1$ if $f \in P_2$ and $\operatorname{Re} w(f) = 0$ if $f \in P_1 \cup P_3$. Thus $|P_0| - |P_2|$ is exactly the real part of the pattern sum. As $s_1 = 2(1 + i)$, $s_2 = 0$, $s_3 = 2(1 - i)$ and $s_4 = 4$, we see immediately that after substitution the real part of each term is 0, so $|P_0| = |P_2|$.

We were first led to our study of the orbits of a permutation group by our desire to count the number of graphs within isomorphism. We realized that this amounted to counting the orbits of the group Γ_n^* acting on $\{0, 1\}^X$, where $X = V^{(2)}$ and Γ_n is the permutation group acting on X which is induced by the symmetric group acting on V. So according to Pólya's theorem our problem is solved when we know the cycle sum of the permutation group Γ_n. It is now a routine matter to write down an explicit expression for this

cycle sum, though we don't display it here since its form is not very inspiring. Furthermore, except for small values of n it is too unwieldy for practical calculations and it is much easier to use asymptotic formulae derived by random graph techniques (see Exercises 22 and 23 of Chapter VII).

We remark finally that an extension of Pólya's theorem covers the case when there is also a group acting on the range of the functions. For instance, if we let S_2 act on $\{0, 1\}$ in the example above, we do not distinguish between a graph and its complement, and may thereby compute the number of graphs which are isomorphic to their complements.

Exercises

1. Draw the Cayley diagram of the quaternion group $\langle a, b \mid a^2 = b^2 = (ab)^2 \rangle$.

2. Find the orders of $\langle a, b \mid a^3 = b^3 = 1, ab = ba \rangle$, $\langle a, b \mid a^4 = b^3 = 1, ab = ba \rangle$ and $\langle a, b \mid a^3 = b^4 = (ab)^2 = 1 \rangle$.

3. Use Euler's formula and information about the Cayley diagram to deduce that in the group $\langle a, b, c \mid a^5 = b^3 = c^2 = (abc)^{-1} \rangle$ we have $a^{610} = 1$.

4. Verify that the Cayley diagram associated with the trefoil knot is the diagram described in Figure VIII.5.

5. Write down presentations of the knots shown in Figure VIII.13. Show that the group of the quinquefoil is isomorphic to $\langle f, g \mid f^5 = g^2 \rangle$ and the group of the tweeny is $\langle a, b \mid ababa^{-1}b^{-1}a^{-1}\, babab^{-1}a^{-1}b^{-1} \rangle$.

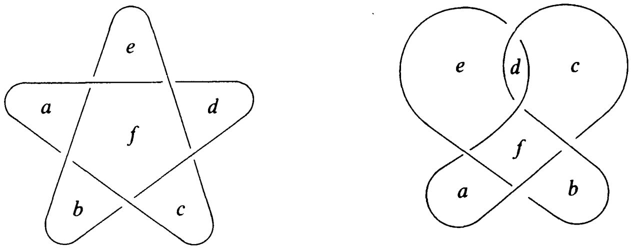


Figure VIII.13. The quinquefoil and the tweeny.

6. Prove that no two of the groups of the knots shown in Figures VIII.4 and VIII.13 are isomorphic.

7. A closed orientable surface of genus 2 is obtained by identifying pairs of non-adjacent sides of an octagon, say as in Figure VIII.14. The fundamental group has a presentation $\langle a_1, a_2, a_3, a_4 \mid a_1 a_2^{-1} a_4 a_1^{-1} a_3 a_2 a_4^{-1} a_3^{-1} \rangle$. Show that the Cayley diagram is a tessellation of the hyperbolic plane. Deduce that any non-empty reduced word W equal to 1 must contain a subword of length at least 5 which is part of the cyclically written relator or its inverse. (A reduced word is one in which no generator occurs next to its inverse.) [*Hint.* Consider the part of the walk W furthest from 1.]

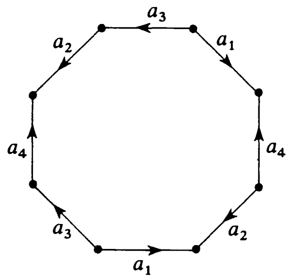


Figure VIII.14. An orientable surface of genus 2.

8. Show that the *dihedral group* D_n, the group of symmetries of a regular n-gon, has a presentation of the form $\langle a, b \mid a^n = b^2 = (ab)^2 = 1\rangle$. What is its Cayley diagram?

9. Give a group whose Cayley diagram is the truncated cube having 8 triangular and 6 octagonal faces.

10. Draw the Cayley diagram of (i) $\langle a, b \mid ab = ba\rangle$ in the Euclidean plane, (ii) $\langle a, b \mid b^n, ab = ba\rangle$ on an infinite cylinder, (iii) $\langle a, b \mid a^m = b^n = 1, ab = ba\rangle$ on a torus, (iv) $\langle a, b, c \mid a^4 = b^4 = c^4 = abc = 1\rangle$ in the hyperbolic plane.

11. Check the examples of Cayley diagrams illustrated in Figure VIII.3 and prove Theorem 1.

12.$^-$ Let $A = \langle a, b \mid a^3 = b^4 = (ba)^2 = 1\rangle$ and $B = \langle b\rangle$. What is the Schreier diagram of A mod B?

13.$^-$ A group A is generated by a, b and c, the Schreier diagram of A modulo a subgroup B is shown in Figure VIII.15. Read off a set of generators of B.

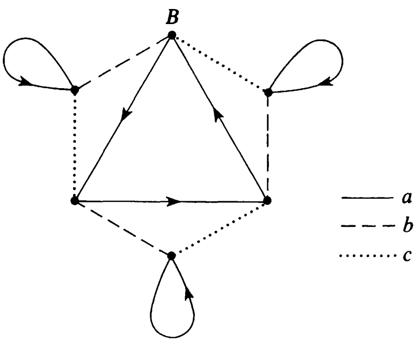


Figure VIII.15. The Schreier diagram of A mod B.

14. Let A be the free group on a, b and c, and let B be the subgroup consisting of all squares. What is the Schreier diagram of A mod B? Find a set of free generators for B.

15. How many subgroups of index 2 are there in a free group of rank k?

16. How many subgroups of index n are there in a free group on 2 generators?

17. Show that a subgroup B of a finitely generated free group F is of finite index in F iff B is finitely generated and there is a natural number n for which $W^n \in B$ for every word W.

18. What is the automorphism group of the Petersen graph (shown in Figure VIII.11)? Find the automorphism group of the Kneser graph $K_s^{(r)}$, where $s \geq 2r + 1$ (see p. 99. Deduce that the automorphism group of $K_{2r+1}^{(r)}$ is 3-*arc-transitive*, that is any path of length 3 can be mapped into any other path of length 3 by an automorphism.

19. The *Tutte* 8-*cage* has vertices n_1, n_3, n_5 for $n = 1, 2, \ldots, 10$, with edges joining n_5 to n_1 and n_3, and n_i to m_i iff $|n - m| \equiv \pm i \pmod{10}$. Show that the automorphism group of the Tutte 8-cage is 5-arc-transitive.

20. Find the automorphism group of the Grötzsch graph (see p. 99).

21. Does every tree have an Abelian automorphism group?

22. Construct a non-trivial tree with trivial automorphism group.

23. Let $f(n)$ be the minimal size of a graph of order with trivial automorphism group. Prove that $f(n) < n$ for every $n \geq 7$ and $f(n)/n \to 1$ as $n \to \infty$.

24. Show that if $\pi \in \operatorname{Aut} G$ has k odd cycles and l even cycles then G has at most $k + 2l$ simple eigenvalues.

25. Show that a connected graph G of odd order whose automorphism group is vertex transitive has exactly one simple eigenvalue.

26.$^-$ Let A be the adjacency matrix of a graph G. Show that $(A^l)_{ij}$ is the number of v_i-v_j walks of length l.

27. Given $k \geq 2$, $p_0(x), p_1(x), \ldots$ be polynomials defined by $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2 - k$ and

$$p_l(x) = xp_{l-1}(x) - (k - 1)p_{l-2}(x), \qquad l \geq 3.$$

Show that if A is the adjacency matrix of a k-regular graph then $(p_l(A))_{ij}$ is the number of v_i-v_j *paths* of length l.

28. Complete the details of the second proof of Theorem 8, as suggested on p. 161.

29. Prove that the matrix J (all of whose entries are 1) is a polynomial in the adjacency matrix A iff G is regular and connected.

30. Combine Theorem 5(vi) and Theorem 1 of Chapter V to deduce that $\chi(G) \leq \lambda_{\max}(G) + 1$.

31. Let $f_G(x) = \sum_{k=0}^{n} c_k x^{n-k}$ be the characteristic polynomial of (the adjacency matrix of) a graph G. Show that $c_0 = 1$, $c_1 = 0$, $c_2 = -e(G)$ and $-c_3$ is twice the number of triangles in G.

32. Let $f_G(x)$ be the characteristic polynomial of G. Show that if $e = xy$ is a bridge of G then

$$f_G(x) = f_{G-e}(x) - f_{G-\{x,y\}}(x).$$

Now let F be a forest with $2n$ vertices and denote by d_k the number of k-element sets of independent edges. [Thus d_n is the number of 1-factors.] Prove that

$$f_G(x) = x^{2n} - d_1 x^{2n-2} + d_2 x^{2n-4} - \cdots + (-1)^n d_n.$$

What are the possible values for d_n?

33. Let G be a connected k-regular graph containing an odd cycle. At time 0 put a counter on a vertex. For each counter that is on a vertex x at time t, place a counter on every vertex adjacent to x at time $t + 1$ and remove the counters from t. Show that $n_t(x)/k^t$ tends to a limit as $t \to \infty$, where $n_t(x)$ is the number of counters on x at time t. What is the corresponding assertion if the counters are not removed from the vertices? [*Hint.* There is an orthonormal basis consisting of the eigenvectors of the adjacency matrix.]

34. A k-regular graph G of order n is such that any two non-adjacent vertices can be mapped into any other such two by some automorphism of G. Show that G is strongly regular. What are the eigenvalues of G? What is the condition that k and n have to satisfy?

35. Let C be the collapsed adjacency matrix of a highly regular graph. Show that $(C^l)_{11}$ is the number of walks of length l from a vertex to itself. Interpret the other entries of C^l.

36. Why must the collapsed adjacency matrix of the Petersen graph, as shown in Figure 11, have rational eigenvalues?

37. In the graph G every two adjacent vertices have exactly one common neighbour and every two non-adjacent vertices have exactly two common neighbours. Show that G is regular of degree k and has order $2k^2 + 1$, for some k in $\{1, 2, 7, 11, 56, 497\}$.

38. Let G be a strongly regular graph of order 100 with parameters $(k, 0, b)$. What are the possible values of k? Find the eigenvalues when $k = 22$.

39. Make use of the result of Exercise 19 to calculate the eigenvalues of the Tutte 8-cage.

40. Give detailed proofs of Theorem 10 and Corollary 11.

41. Prove that the number of trees with $n - 1 \geq 2$ labelled *edges* is n^{n-3}.

42. Show that a given vertex has degree 1 in about $1/e$ of all labelled trees, where $e = 2.71828\ldots$. (Cf. Exercise 7 in Chapter VII.)

43. Denote by $R(n, k)$ the number of trees with n distinguishable vertices of which exactly k have degree 1. Prove that

$$\frac{k}{n} R(n, k) = (n - k)R(n - 1, k - 1) + kR(n - 1, k).$$

44. Prove that there are

$$\frac{1}{m!} \sum_{j=0}^{m} (-\tfrac{1}{2})^j \binom{m}{j} \binom{n-1}{m+j-1} n^{n-m-j} (m+j)!$$

acyclic graphs on n distinguishable vertices have $n - m$ edges. Deduce that there are $\frac{1}{2}(n - 1)(n + 6)n^{n-4}$ forests with two components.

45. Show that the cycle sum of the symmetric group S_n acting on the usual n letters is

$$\sum \frac{n!}{\prod_{k=1}^{n} k^{j_k} j_k!} a_1^{j_1} a_2^{j_2} \cdots a_n^{j_n},$$

where the summation is over all partitions $j_1 + 2j_2 + \cdots + nj_n = n$.

46. What is the cycle sum of S_5 acting on the unordered pairs of 5 elements? How many non-isomorphic graphs are there on 5 vertices? How many of them have 5 edges?

47. Let $Z(\Gamma, a_1, a_2, \ldots, a_d)$ be the cycle index of a permutation group Γ acting on a set D. Consider the action of Γ on the set of all k-subsets of D. How many orbits are there?

48. Determine the cycle index of the rotations of the cube acting on (i) the vertices, (ii) the edges, (iii) the faces, (iv) the faces and vertices. In how many distinct ways can you colour the vertices using some of n colours? The edges? The faces? The vertices and faces?

49. Show that the cycle sum of C_n, the cyclic group of order n, is $\tilde{Z}(C_n; a_1, \ldots, a_n) = \sum_{k/n} \phi(k) a_k{}^{n/k}$, where $\phi(k)$ is the Euler function.

50. Prove that the cycle index of the dihedral group D_n (cf. Exercise 8) is

$$Z(D_n; a_1, \ldots, a_{2n}) = \tfrac{1}{2}Z(C_n; a_1, \ldots, a_n) + \begin{cases} \tfrac{1}{2}a_1 a_2^{(n-1)/2} & \text{if } n \text{ is odd,} \\ \tfrac{1}{4}(a_2^{n/2} + a_1^2 a_2^{(n-2)/2}) & \text{if } n \text{ is even.} \end{cases}$$

How many bracelets are there with 20 beads coloured red, blue and green?

51. How many distinct ways are there of colouring the faces of a dodecahedron with red, blue and green, using each colour at least once?

Notes

There is a vast literature concerned with group presentations, including the use of Cayley and Schreier diagrams. The basic book is perhaps W. Magnus, A. Karrass and D. Solitar, *Combinatorial Group Theory*, Second Edition, Dover, New York, 1976; the connections with geometry are emphasised in H. S. M. Coxeter, Regular Complex Polytopes, *Cambridge University Press*, New York, 1974. Numerous articles deal with the computational aspect, in particular J. A. Todd and H. S. M. Coxeter, A practical method for enumerating cosets of a finite abstract group, *Proc. Edinburgh Math. Soc.* (2) **5** (1936) 26–34, which was the first paper in this line and J. Leech, Computer proof of relations in groups, in *Topics in Group Theory and Computation* (M. P. J. Curran, ed.), Academic Press, New York, 1977, in which some more recent developments are described.

Max Dehn posed the word problem in Über unendliche diskontinuierliche Gruppen, *Math. Ann.* **71** (1911) 116–144, and gave the above discussed presentation of the group of the trefoil in Über die Topologie des dreidimensionalen Raumes, *Math. Ann.* **69** (1910) 137–168. The word problem was shown to be intrisically connected to logic by P. S. Novikov, On the algorithmic unsolvability of the word problem, *Amer. Math. Soc. Transl.* (2) **9** (1958) 1–22 and G. Higman, Subgroups of finitely presented groups, *Proc. Royal Soc.* A, **262** (1961) 455–475. For the properties of knots and their groups the reader is referred to R. H. Crowell and R. H. Fox, *Introduction to Knot Theory*, Graduate Texts in Mathematics, Vol. 57, Springer-Verlag, New York, 1977.

An exposition of matrix methods in graph theory can be found in N. Biggs, *Algebraic Graph Theory*, Cambridge University Press, New York, 1974. The first striking result obtained in this way, Theorem 9, is due to A. J. Hoffman and R. R. Singleton, On Moore graphs with diameters 2 and 3, *IBM J. Res. Dev.* **4** (1960) 497–504. The connection between graphs and sporadic simple groups is discussed in detail in R. Brauer and C. H. Sah (editors), *Theory of Finite Groups: a Symposium*, Benjamin, Menlo Park, 1969. Since then however several more sporadic groups have been found.

The fundamental enumeration theorem of G, Pólya appeared in Kombinatorische Anzahlbestimmungen für Gruppen und chemische Verbindungen, *Acta Math.* **68** (1937) 145–254. Many enumeration techniques were anticipated by J. H. Redfield, The theory of group-reduced distributions, *Amer. J. Math.* **49** (1927) 433–455. The standard reference book for Pólya-type enumeration is F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, New York, 1973.

Subject Index

Index of Symbols

Graduate Texts in Mathematics

Soft and hard cover editions are available for each volume up to vol. 14, hard cover only from Vol. 15

1 Takeuti/Zaring. Introduction to Axiomatic Set Theory.
2 Oxtoby. Measure and Category.
3 Schaeffer. Topological Vector Spaces.
4 Hilton/Stammbach. A Course in Homological Algebra.
5 MacLane. Categories for the Working Mathematician.
6 Hughes/Piper. Projective Planes.
7 Serre. A course in Arithmetic.
8 Takeuti/Zaring. Axiomatic Set Theory.
9 Humphreys. Introduction to Lie Algebras and Representation Theory. 2nd printing, revised.
10 Cohen. A Course in Simple Homotopy Theory.
11 Conway. Functions of One Complex Variable. 2nd ed.
12 Beals. Advanced Mathematical Analysis, xi, 230 pages. 1973.
13 Anderson/Fuller. Rings and Categories of Modules.
14 Golubitsky/Guillemin. Stable Mappings and Their Singularities.
15 Berberian. Lectures in Functional Analysis and Operator Theory.
16 Winter. The Structure of Fields.
17 Rosenblatt. Random Processes. 2nd ed.
18 Halmos. Measure Theory.
19 Halmos. A Hilbert Space Problem Book.
20 Husemoller. Fibre Bundles. 2nd ed.
21 Humphreys. Linear Algebraic Groups.
22 Barnes/Mack. An Algebraic Introduction to Mathematical Logic.
23 Greub. Linear Algebra. 4th ed.
24 Holmes. Geometric Functional Analysis and Its Applications.
25 Hewitt/Stromberg. Real and Abstract Analysis. 4th printing.
26 Manes. Algebraic Theories.
27 Kelley. General Topology.
28 Zariski/Samuel. Commutative Algebra I.
29 Zariski/Samuel. Commutative Algebra II.
30 Jacobson. Lectures in Abstract Algebra I: Basic Concepts.
31 Jacobson. Lectures in Abstract Algebra II: Linear Algebra.
32 Jacobson. Lectures in Abstract Algebra III: Theory of Fields and Galois Theory.
33 Hirsch. Differential Topology.
34 Spitzer. Principles of Random Walk. 2nd ed.
35 Wermer. Banach Algebras and Several Complex Variables. 2nd ed.
36 Kelley/Namioka. Linear Topological Spaces.
37 Monk. Mathematical Logic.

38 Grauert/Fritzsche. Several Complex Variables.
39 Arveson. An Invitation to C^*-Algebras.
40 Kemeny/Snell/Knapp. Denumerable Markov Chains. 2nd ed.
41 Apostol. Modular Functions and Dirichlet Series in Number Theory.
42 Seere. Linear Representations of Finite Groups.
43 Gillman/Jerison. Rings of Continuous Functions.
44 Kendig. Elementary Algebraic Geometry.
45 Loeve. Probability Theory. 4th ed. Vol. 1.
46 Loeve. Probability Theory. 4th ed. Vol. 2.
47 Moise. Geometric Topology in Dimensions 2 and 3.
48 Sachs/Wu. General Relativity for Mathematicians.
49 Gruenberg/Weir. Linear Geometry. 2nd ed.
50 Edwards. Fermat's Last Theorem.
51 Klingenberg. A Course in Differential Geometry.
52 Hartshorne. Algebraic Geometry.
53 Manin. A Course in Mathematical Logic.
54 Graver/Watkins. Combinatorics with Emphasis on the Theory of Graphs.
55 Brown/Pearcy. Introduction to Operator Theory. Vol. 1: Elements of Functional Analysis.
56 Massey. Algebraic Topology: An Introduction.
57 Crowell/Fox. Introduction to Knot Theory.
58 Koblitz. p-adic Numbers, p-adic Analysis, and Zeta-Functions.
59 Lang. Cyclotomic Fields.
60 Arnold. Mathematical Methods in Classical Mechanics.
61 Whitehead. Elements of Homotopy Theory.
62 Kargapolov/Merzljakov. Fundamentals of the Theory of Groups.
63 Bollobas. Graph Theory—An Introductory Course.
64 Edwards. Fourier Series, Volume 1. second edition.
65 Wells. Differential Analysis on Complex Manifolds.
66 Waterhouse. Introduction to Affine Group Schemes.
67 Serre. Local Fields.
68 Weidmann. Linear Operators in Hilbert Spaces.